Clinical Practice and Cancer Biology	Biochemistry	Physics & Mathematics	Engineering & Information Systems
David B. Agus (USC)	Jonathan Katz (USC)	Vitorio Cristini (UNM)	Anand Asthagiri (NEU)
Dean Felsher (Stanford)	Parag Mallick (SU/USC)	Sam Gambhir (Stanford)	Richard Bonneau (NYU)
Sam Gambhir (Stanford)	Gary Nolan (Stanford)	Murray Gell-Mann (AMI)	W. Daniel Hillis (USC/AMI)
Mitchell Gross (USC)		Paul Macklin (USC)	Carl Kesselman (USC/ISI)
Joshua LaBaer (ASU)		Matteo Pellegrini (UCLA)	Parag Mallck (USC)
Scott Lowe (CSHL)		Dan Ruderman (USC)	Shan Wang (Stanford)
		Tom Tombrello (Cal-Tech/AMI)	

Multiscale Complex Systems Transdisciplinary Analysis of Response to Therapy

Parag Mallick paragm@stanford.edu



Program says - 'Omics Biomarker'

Reality: 'Omics,' 'Biomarkers'



Introduction to the USC PSOC

Developing models of cellular regulation

The complicated relationship between the tumor proteome and the circulating proteome.



Our Overarching Goal

The main theme of our PSOC is to develop a multi-scale virtual cancer model that is able to accurately predict the growth and response to therapy of a tumor given a set of measured inputs.





Scales in Biology & Physics

Biologist describes across scales

Scale	Biological description	
Genomic	Genes, transcripts, proteins	Small scale
Organelle	Protein-protein interactions	
Cell	Pathways, phenotypes	
Tissue	Inter-cellular signaling	V
Organ	Physiological processes	Large scale
Organism	Health, disease	, C
Population	Epidemiology	

Physicists model across scales

Scale	Physical model	
Sub-nuclear	Standard model	
Atomic	Quantum mechanics	
Molecular	Chemistry	
Mesoscopic	Statistical and condensed matter physics	
Macroscopic	Classical physics	
Astronomical	Cosmology	



Resistance is a multi-scale problem

The Environment or Host

 Drug never hits the target, physical blockades exist, hypoxia or other mechanical variation affects drug effectiveness.

The Target

 Something about the target is 'broken'

Downstream of the target

 A cell's response circuitry is broken or something is compensating



We believe an integrative, multi-scale approach is necessary to develop accurate, useable models to study therapeutic response in cancer. In particular, we note that subtle molecular-scale perturbations (e.g. mutation in a gene) can produce dramatic, tumor-scale (e.g. invasiveness) and organism scale (e.g. responsiveness to therapy) affects.



Our Specific Biological System

- Burkitt's Lymphoma, Eµ-Myc Model
- Recapitulates typical genetic and pathological features of human Non-۰ Hodgkin's lymphomas
- Tumors arise with relatively short latency and high penetrance ۲
- Therapy is performed in immuno-competent mice
- Lymphoma cells can be cultured and transplanted into syngenic, non-• transgenic recipient mice.
- The same cells can be studied *in vitro* and *in vivo* for cross-scale integration



In Vivo Treatment Response

Model Structure





SPR experiments protein binding dynamics

9



Using Multi-scale Systems Approaches to Uncover Biomarkers and Mechanisms

Topics

Background and Overview USC PSOC

Modeling Cellular Regulation Transcript-level Upscaling to Protein Connecting Protein and Phenotype

Quantitative models of the relationship between the tumor and circulating proteomes to aid biomarker discovery

Other Random Fun. Cell Mechanics (w/ Scott Manalis)



Research Project 1 Multi-Regulatory Scale Models of Cellular Dynamics





11

The Molecular/Cellular Team



Overview of Project

Develop a computational model that operates at and below the cellular scale and across multiple time scales to describe how the genetic background and chemical/environmental context of a cell regulate its behavior and engender phenotypes (e.g. response to therapy) that ultimately impact the tumor and host.







Project Overview

Problem: Given a set of cell-intrinsic (genotype) and cell-extrinsic inputs (environment) and some calibrants (e.g. transcript, protein levels) infer a cell's resulting state and state-evolution function (phenotype).

Strategy: Rigorously measure a large number of molecular and cellular parameters in steady state, and in response to diverse perturbations. Use those measurements with diverse data to build our model. Simulate perturbations. Validate with additional experiments.

Deliverable: A computational model that describes cell regulatory dynamics at multiple scales, and 'clicks' into the tumor-scale model.



Drug Resistance





Drug Resistant

Drug Resistance





Questions

- What are the processes going on inside cells that govern how they will respond to therapeutic (or other) perturbation?
- Can we describe those processes and possibly predict how novel interventions will act?
- When cells are dying (or not dying) what does that process entail?
- Can't we just do RNASEQ and call it good?



From Measurements to State



The Geneticist's Approach





Hmmm...but that doesn't work all the time...

Table 1. Phenomena complicating the concept of the gene

Phenomenon	omenon Description	
Gene location and structure		
Intronic genes	A gene exists within an intron of another (Henikoff et al. 1986)	Two genes in the same locus
Genes with overlapping reading frames	A DNA region may code for two different protein products in different reading frames (Contreras et al. 1977)	No one-to-one correspondence between DNA and protein sequence
Enhancers, silencers	Distant regulatory elements (Spílianakis et al. 2005)	DNA sequences determining expression can be widely separated from one another in genome. Many-to-many relationship between genes and their enhancers.
Structural variation		
Mobile elements	Genetic element appears in new locations over generations (McClintock 1948)	A genetic element may be not constant in its location
Gene rearrangements/structural variants	DNA rearrangement or splicing in somatic cells results in many alternative gene products (Early et al. 1980)	Gene structure is not hereditary, or structure may differ across individuals or cells/tissues
Copy-number variants	Copy number of genes/regulatory elements may differ between individuals (lafrate et al. 2004; Sebat et al. 2004; Tuzun et al. 2005)	Genetic elements may differ in their number
Epigenetics and chromosome structure		
Épigenetic modifications, imprinting	Inherited information may not be DNA-sequence based (e.g., Dobrovic et al. 1988); a gene's expression depends on whether it is of paternal or maternal origin (Sager and Kitchin 1975)	Phenotype is not determined strictly by genotype
Effect of chromatin structure	Chromatin structure, which does influence gene expression, only loosely associated with particular DNA sequences (Paul 1972)	Gene expression depends on packing of DNA. DNA sequence is not enough to predict gene product.



Genome Res. 2007 17: 669-681

The Geneticist's Approach





A simple model for how cells work

Cell Phenotype is controlled by 'Gene X'





Flow of Biological Information





Different Models of Cellular Regulation



Overall Approach

1. Inferelator Magic to Derive Transcriptional Regulatory Network





2. Glue Transcript to







Analysis "Standard Workflow"





Starting Data

Species	Normal	Lymphoma	Outgroup	Total
Mouse	688	295	41	1024
Human	445	447	53	1025

- Goals
 - Identify state-specific functional modules
 - Search for differential expression over known time course
- Applied pipeline to public microarray data
 - Multi-species biclustering
 - Inference on biclusters
- Follow-up analysis
- Identification of state-specific biclusters
 - Analysis of connected state-specific groups
 - Overlays of project-generated time course data







- Eµ-Myc/p53-/- (resistant)
- Eµ-Myc/pArf-/- (sensitive)
- Drug doses based on patient serum levels (Cornelius/Lowe lab)





Full Transcriptomic Network



Maf treatment of p53-/- vs arf-/-





Predicting Network Perturbations





Overall Approach

1. Inferelator Magic to Derive Transcriptional Regulatory Network





2. Glue Transcript to







Comparing Changes in Protein and Transcript Abundances

ARF-/- 24H : 0H Post Mafosfamide Treatment



Log (Relative Transcript Abundance)

No obvious correlation between transcript and protein changes in abundance



Protein and mRNA Changes Over time Between Eμ-Myc p53^{-/-} & Arf^{-/-} Lines



> Involved in DSB repair.



Part of Nolan Lab Panel



>Inhibits tumor promoter-induced neoplastic transformation.



Involved in cell cycle progression.



Multi-Scale Regulatory Model



Predicting Protein Levels from Transcript Levels


Time Course Protein Level Prediction

Arpc3



- (Early) Summary
 - Published
 degradation rates
 may be low
 - Constant rates
 from short time
 course
 inadequate
 - Does not account
 for regulation of
 degradation (*e.g.*,
 ubiquitinylation)



Overall Approach

1. Inferelator Magic to Derive Transcriptional Regulatory Network





2. Glue Transcript to







Mass Cytometry versus Fluorescence

Fluorescence

- Up to 12 colors can be "routine"
- 17 colors have been reported
- High background

Elemental Mass Spectrometry

- Up to 100 non-biological elemental mass channels
- No compensation required
- Dynamic range 10⁴
- No autofluorescence







PTM-one Analysis Challenges

Single-cell vs. gmish Small number of components

A LOT of data and slices

Completely different sample preparation



CyTOF Panel	Function
p27	Cell Cycle
p21	Cell Cycle
Cyclin B1	Cell Cycle
p-Histone H3 (pS28)	Cell Cycle
p-CDK1 (Y15)	Cell Cycle
p-CHK1 (S345)	Cell Cycle/ Checkpoint
p-Chk2 (pT68)	Cell Cycle/ Checkpoint
p-pRb (S807/811)	Cell Cycle/ Proliferation/ Apoptosis
p-H2AX (S139)	DDR
p-ATM (pS1981)	DDR
p-BRCA1 (S988)	DDR
p-53BP1 (S1778)	DDR
PAR	DDR
p-p53 (S37)	DDR/ Apoptosis
p-p53 (S15)	DDR/ Apoptosis
cleaved-Caspase3	Apoptosis
cleaved-PARP	Apoptosis
p-Bcl-2 (S70)	Survival
XIAP	Survival
McI1	Survival
p-AMPK (T172)	Metabolism
p-S6 (pS235/36)	Protein Translation
p-Creb (pS133)	Transcription
mCD90	Surface Marker
B220	Surface Marker



Spanning Tree Progression Analysis of Density-Normalized Events (SPADE) Trees (Total Cell Numbers)



Spanning Tree Progression Analysis of Density-Normalized Events (SPADE) Trees (Total Cell Numbers)



Spanning Tree Progression Analysis of Density-Normalized Events (SPADE) Trees (Total Cell Numbers)





Spanning Tree Progression Analysis of Density-Normalized Events (SPADE) Trees (p-H2AX)





Spanning Tree Progression Analysis of Density-Normalized Events (SPADE) Trees (p-H2AX)





Overall Approach

1. Inferelator Magic to Derive Transcriptional Regulatory Network





2. Glue Transcript to







Prediction From Perturbation



Prediction From Perturbation



State-Relevant Sub-Network





Cell State-Specific Expression







Cell State-Specific Expression

-0.52



0.0

0.85



Subnetwork Expression to Cell State

- Hypothesis: each cell state has a characteristic expression pattern over sub-network
- Use patterns to train statistical model of cell state
 - Input: average bicluster expression
 - Output: vector of cell state probabilities



Classification By Bicluster Expression







Using Multi-scale Systems Approaches to Uncover Biomarkers and Mechanisms

Topics

Background and Overview USC PSOC

Modeling Cellular Regulation Transcript-level Upscaling to Protein Connecting Protein and Phenotype

Quantitative models of the relationship between the tumor and circulating proteomes to aid biomarker discovery

Other Random Fun.

Cell Mechanics (w/ Scott Manalis)



The Overall Objective



Does the patient have cancer?

- Is that cancer aggressive/invasive?
- Is the cancer likely to respond to drug X?
- Is the cancer actually responding to drug X?



The unbiased discovery approach





The Biomarker Discovery Problem (Waldo version)



Cancer Patients



Healthy Controls



A (slightly) More Realistic Example... Find the Differences – part deux

Cancer Patients

Healthy Controls



Easier Challenge – look in a relevant place – *THEN* find the differences

Cancer Patients





Healthy Controls





Rethinking the problem - Where do biomarkers come from?



Hypothesis: A subset of markers are derived from the tumor...?

(and also host response, which we hypothesize to be somehow related to the tumor)

Biomarkers are host-scale measurements that tell us about tumor and cell-scale phenomena



Key Questions of the Biomarker Discovery Process



Hidden Assumptions

• There is a signal to be found in the tissue.



• That signal makes it from the tumor into the circulation



Marker Discovery Problem Re-Statement

Step 1: Identify proteins that are indicative of the aberrant state/trajectory of cancer cells (or perhaps their environment)

- Question: Are there any?
- **Step 2:** Characterize the composition of the tumor.
 - Question: what are the evolutionary forces at work?

Step 3: Identify CIRCULATING/CELL-SURFACE proteins indicative of the presence and state/trajectory of a tumor

- Question: How do these relate to the proteins in Step 1?







Data Collection





Relationship between Tumor levels and Circulating Levels



Log Fold Change in Tumor



Explanations

- Non-Uniform processes of transfer from tumor to circulation.
- Background Levels



Summary of tumor proteins identified and quantified in each experiment and their cellular location

Xenograft mouse model	Average size of tumors	Tissue type	Proteins identified	Cellular location			
				Extracellular	Non-	Not	
					extracellular	annotated	
A431 small	750 mm ³	plasma	103	42	54	7	
		tumor	2314	170	1882	262	
A431 large	1300mm ³	plasma	87	38	42	7	
		tumor	2099	163	1705	231	

In addition, 450 and 499 mouse proteins were identified in A431s and A431l plasma respectively.



MA plots of A431 small and A431 large human & mouse peptides in plasma



Bigger tumor leads to more proteins detected in plasma, but not in 1:1 association







Protein abundance vs. tumor proteins observed /not observed in plasma

Percentage of tumor proteins observed in plasma by spectral count





	Observed in plasma			Not observed in plasma				Chi-square test	
instability index score*	<= 25%	25-50%	50-75%	>75%	<= 25%	25-50%	50-75%	>75%	
A431s	37	20	12	11	542	558	566	568	7.3e-05
A431r	32	22	12	11	493	503	512	514	0.001
	Stable Unstable			e Stable	9	Ur	stable	9	

* The higher the instability index score, the lower the protein stability is.



	multi	variate	marginal		
	Coefficie nt	P value	Coefficient	P value	
In extracellular	1.95	7.1e-13	1.91	1.2e-13	
Stability	0.87	0.001	1.06	9.8e-06	
Spectral counts	0.44	9.4e-13	0.43	3.8e-15	
# of tryptic peptides	-0.008	0.05	-0.0008	0.82	


Ability to Predict Circulating Tumor Derived Proteins



1-Specificity



Summary

- We are working to develop approaches for modeling cell behaviors and identifying the genes/proteins that are most impacting cell-states affiliated with DDR
- 2) Cellular control systems operate at multiple scales (transcript, protein, PTM...)
- 3) In our system there are clearly multiple stages of cellular response – Damage Sensing, Damage Response and then Several phases of cell death that stall differently in different cells



Acknowledgements





Using Multi-scale Systems Approaches to Uncover Biomarkers and Mechanisms

Topics

Background and Overview USC PSOC

Modeling Cellular Regulation Transcript-level Upscaling to Protein Connecting Protein and Phenotype

Quantitative models of the relationship between the tumor and circulating proteomes to aid biomarker discovery

Other Random Fun. Cell Mechanics (w/ Scott Manalis)



Biomechanics of Metastatic progression



Adapted from Kalluri and Weinberg JCI 2009

Numerous molecular and physical properties change in the process of invasion to a distant site

What can we learn about the tumor and its progression by measuring **deformability** and **friction** of cancer cells in an in vitro system?

Ultimately, could such a system be used to identify and characterize circulating tumor cells?



Cell Mechanics Has a Long History





Pelling and Horton, Plugers Arch- Eur J Physiol (2008) 456:3-12

Methods for Probing Cell Mechanics

Micropipette Aspiration



http://newton.ex.ac.uk/research/biomedicalold/membranes/vesicle.html

Atomic Force Microscopy (AFM)



Bao and Suresh, Nature Materials, 2003

Intracellular Nanorheology (IN)



Panorchan et al. Methods in Cell Biology 2007

- Source of many of our classical models of global cellular deformation
- Problems: low throughput, irreproducible, low accuracy

- High accuracy
- Good for measuring membrane properties
- Mainly used to study local deformation

- Infer material properties from Brownian motion
- Often limited to viscosity measurements



Microchannels for Cell Mechanics



Gabriele et al, 2010, Lab Chip

- Higher throughput than micropipettes
- But still relies on optical methods to measure cell size/trajectory – imprecise



A new approach to measure biomechanics in high throughput

 New <u>Suspended Microchannel Resonator</u> approach is rapid, precise and can measure both cell rigidity (squishyness) and cell-surface friction (slimyness)









High metastatic cancer cells are squishier and in some cases slimier



Mouse model with single transcription factor addition Tmet vs. Tmet-Nkx2

Mouse model of lung cancer Tmet vs. Tnonmet

S. Byun, S. Son, D. Amodei, N. Cermak, J. Shaw, M. Winslow, T. Jacks, P. Mallick and S. Manalis. Characterizing deformability and surface friction of cancer cells, PNAS, *in revision*.

Human lung cancer cell lines Mesenchymal (H1975) vs Epithelial (HCC827)



Acknowledgements

