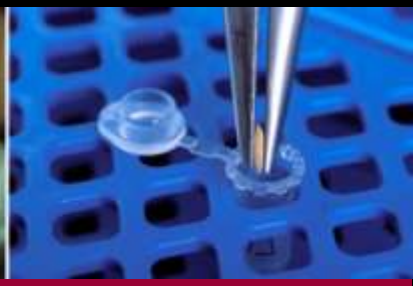
ASU BIODESIGN INSTITUTE
ARIZONA STATE UNIVERSITY

SAIC From Science to Solutions
SAIC-Frederick, Inc.
A subsidiary of Science Applications International Corporation

Diagnostic Biomarkers and the Human Protein-Drug Interaction Network: Implications and Opportunities for Targeted Oncology Therapies

Presentation at SAIC-Frederick
1 May 2008

Dr. George Poste, Director
The Biodesign Institute, Arizona State University
Tel: 480-727-8662, e.mail: george.poste@asu.edu
www.biodesign.asu.edu



ASU PROPRIETARY

NCI's Vision for Cancer and the Nation's Healthcare System

An interoperable "system" of evidence based molecular oncology where disease is either prevented or detected early and treated on an evidence-based and personalized basis

Lessons Learned

- molecular biology and reductionism

  to

  systems biology and the design of biological networks
- insular focus on single analytical platforms

  to

  integrated fusion of datasets from diverse platforms
- extravagant waste of under-powered studies of biomarkers and Rx investigational trials

  to

  rigorous validation standards and robust statistical algorithms for biomarker verification and validation and Rx efficacy
- one drug : one target strategies for Rx screening and translation

  to

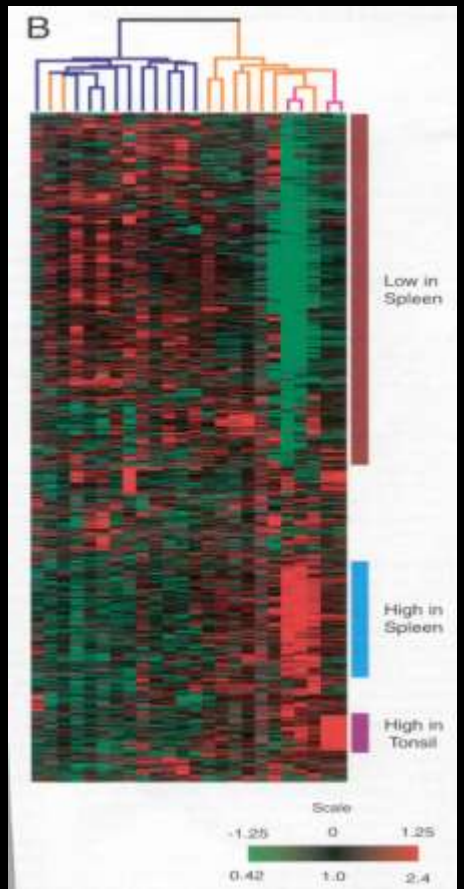
  drug-target network analysis and importance of Rx promiscuity in acting on linked modules in biological pathways

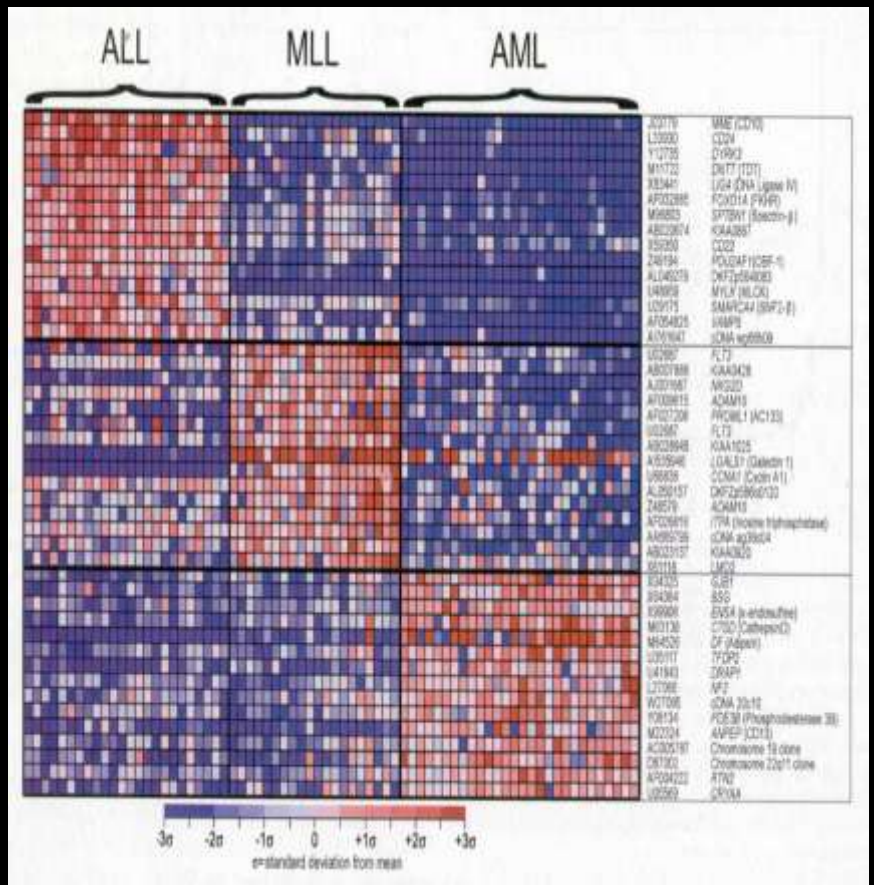
The Intellectual Foundation of Molecular Medicine and Rational Therapeutics

- mapping network perturbations in major diseases
  - cellular specificity and network design
  - disease-associated 'biosignatures' of network disruption
  - 'signatures' for disease predisposition
- network pharmacology
  - design of targeted Rx
  - customized selection of optimum Rx
- technical foundation for progressive evolution of targeted therapeutics and personalized medicine

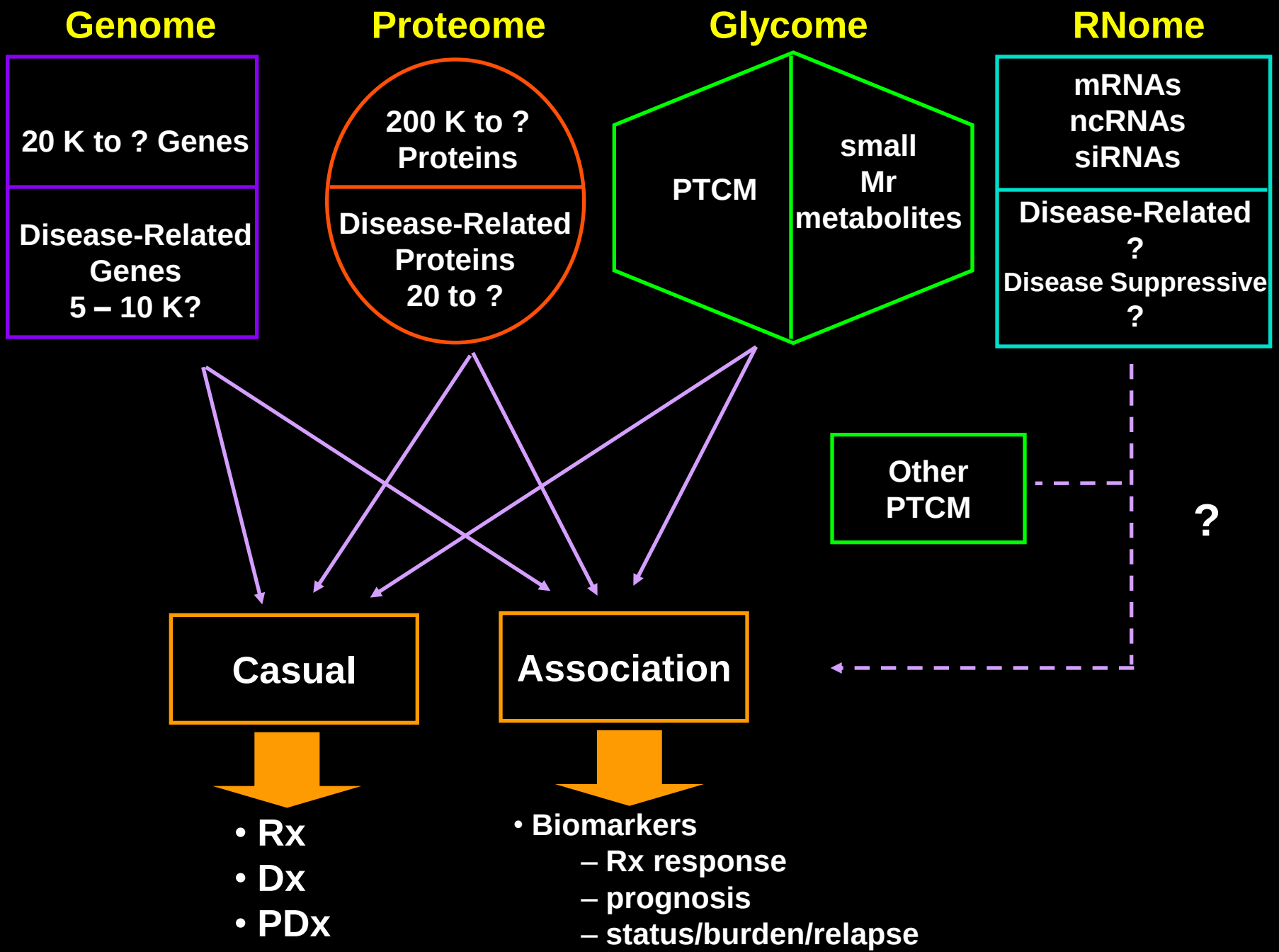
Molecular Diagnostics and Biomarkers

The Fundamental Technology Platforms For Molecular Medicine and Vital Elements of the Future Healthcare Value Chain

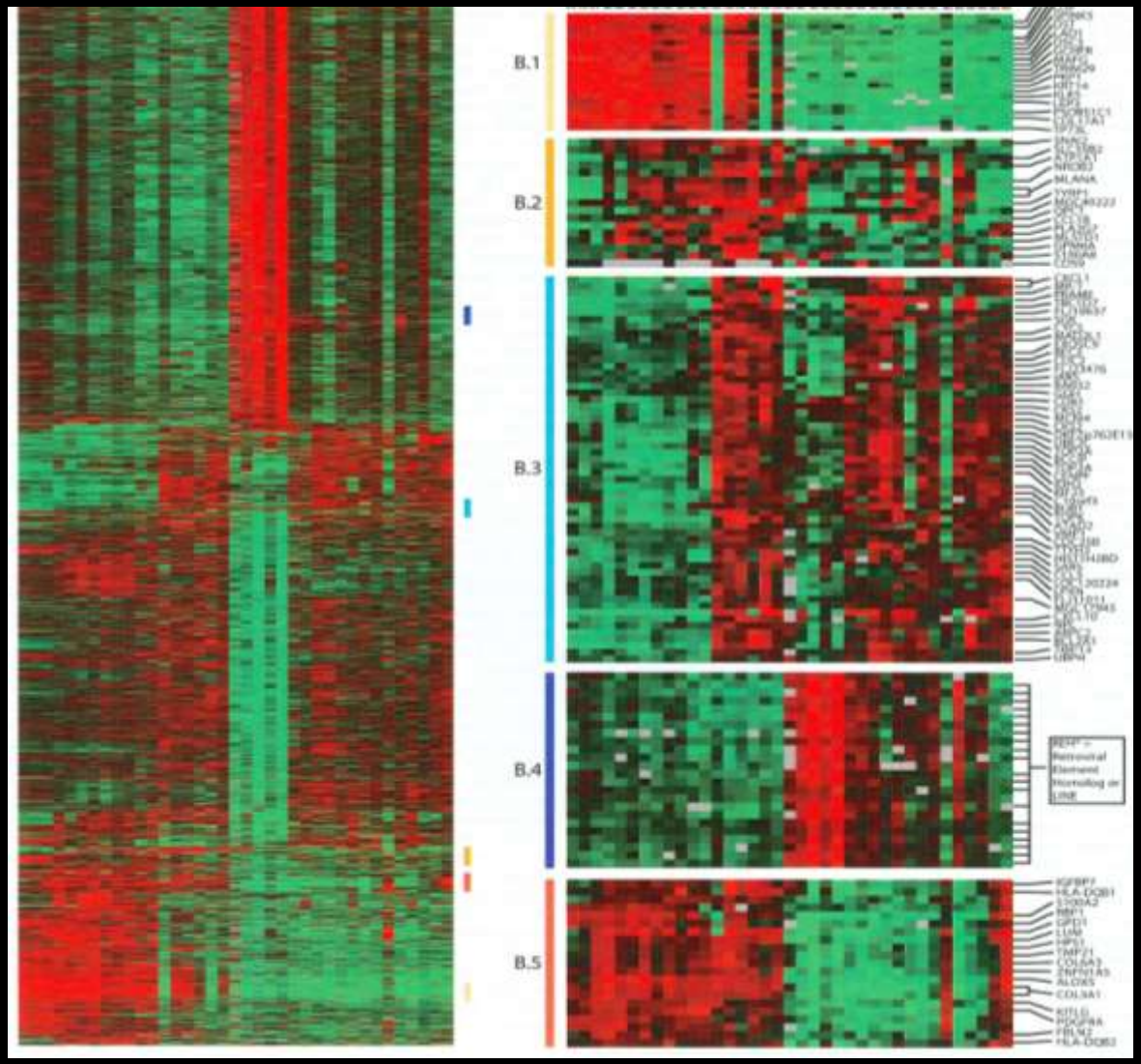




Identification of Disease-Associated Biomarkers



Disease Subtyping: Next-Generation Molecular Diagnostics (MDx) and A New Molecular Taxonomy of Disease

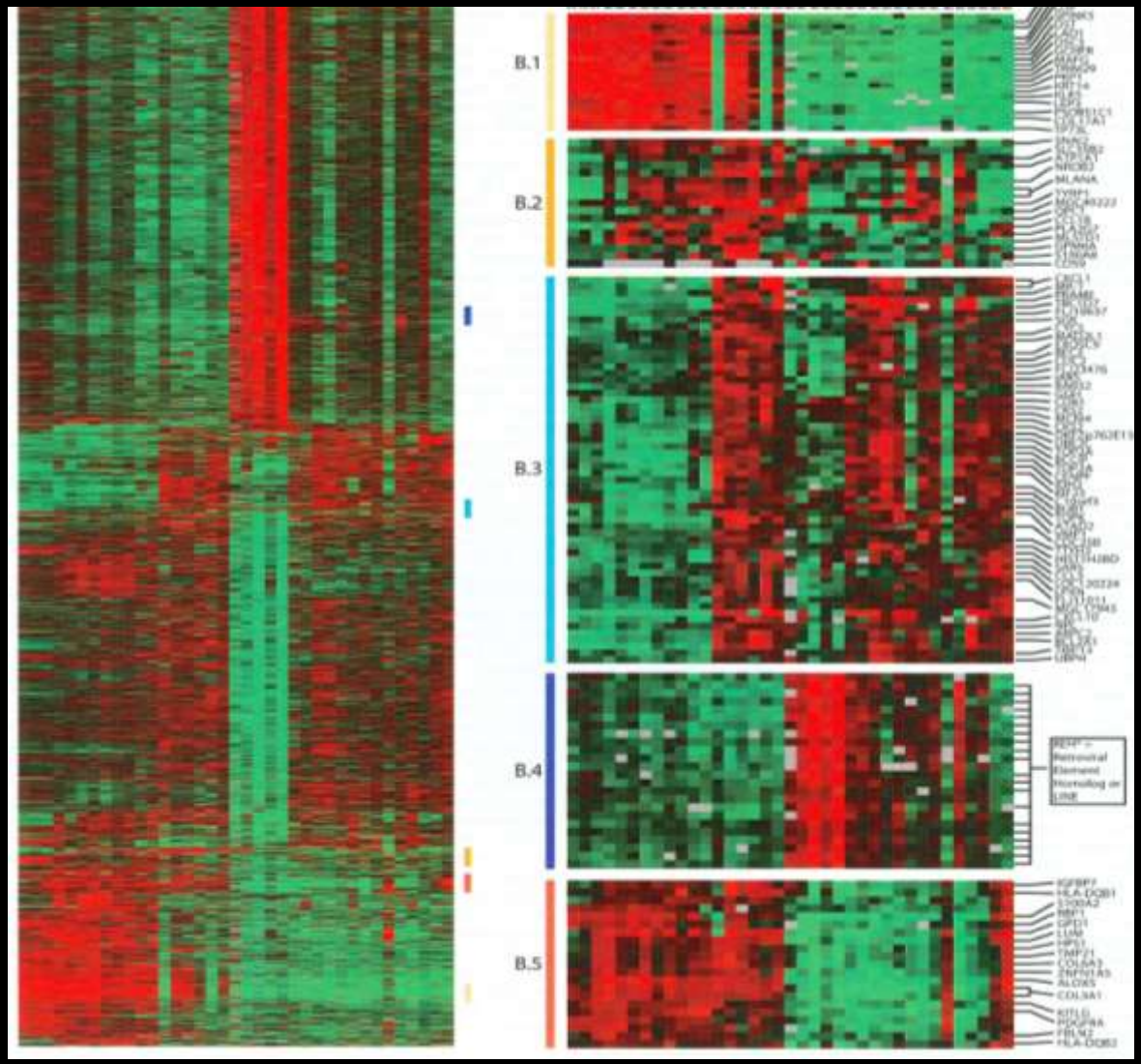


Dx Platforms

- massive parallelism
- miniaturization
- automation
- rapid
- POC

B1 skin, B2, melanocytes, B3, melanoma, B4 and 5 metastatic melanoma
From: C. Haqq et al. (2005) 102, 6092

Disease Subtyping: Next-Generation Molecular Diagnostics (MDx) and A New Molecular Taxonomy of Disease



Dx Platforms

- massive parallelism
- miniaturization
- automation
- rapid
- POC

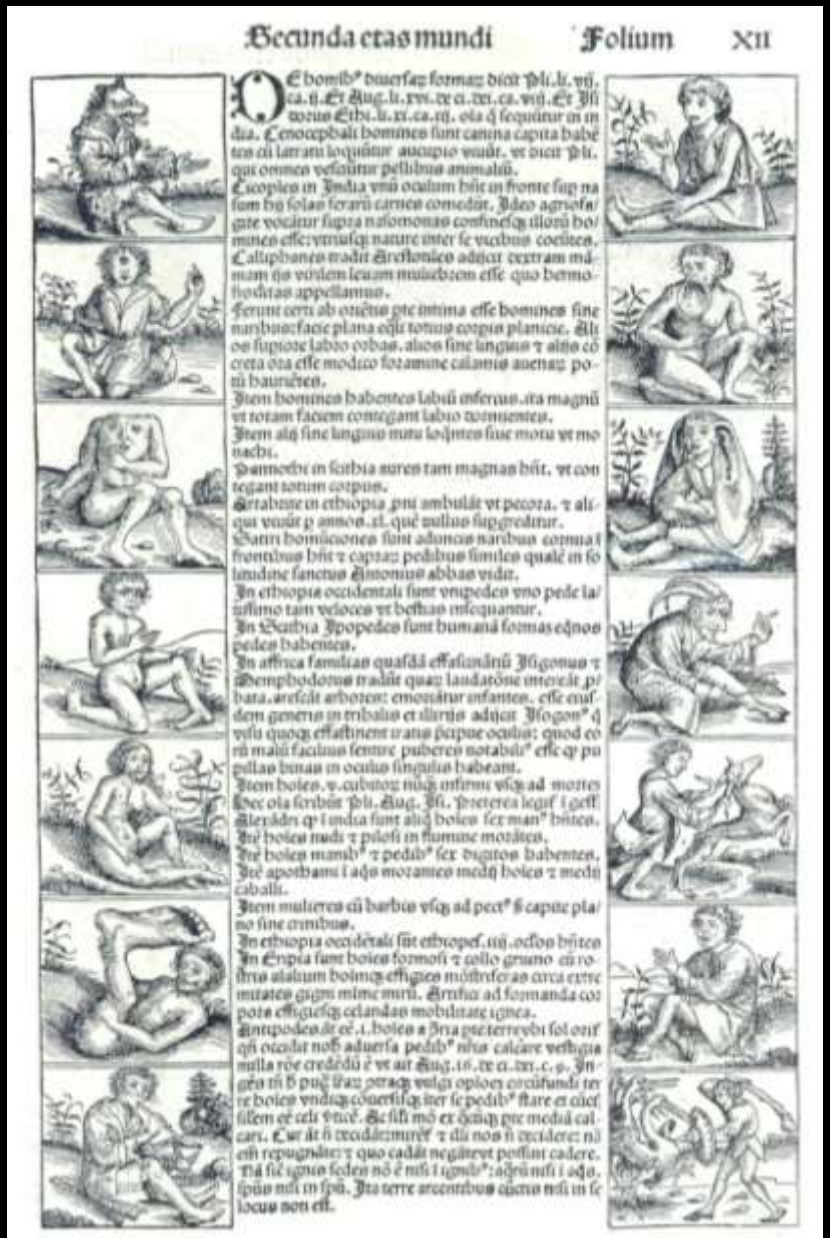
RIGHT Rx for RIGHT DISEASE SUBTYPE

B1 skin, B2, melanocytes, B3, melanoma, B4 and 5 metastatic melanoma
From: C. Haqq et al. (2005) 102, 6092

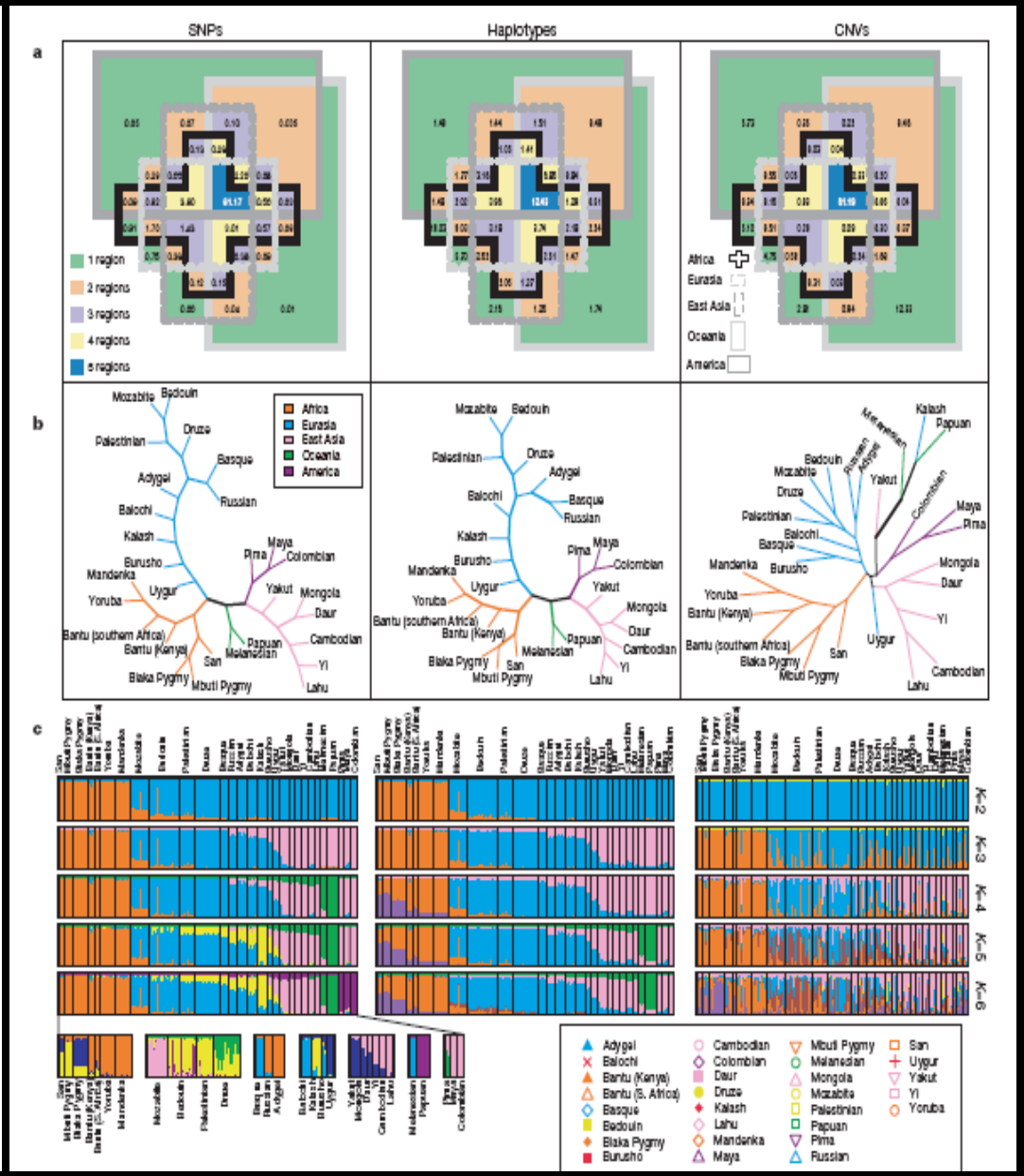
Predictive Gene Lists (PGLs) of Altered Gene Expression in Cancer as Diagnostic and Prognostic Tests

- wide variation in PGLs reported in different publications for claimed similar patients
  - minimal overlap between studies
- discriminatory power of classifiers not reproducible when tested on published cross-sets of samples
- trigger for new regulatory oversight and sophisticated validation protocols
  - IVDMIAs (FDA)

Supposed Variations in Human Anatomy
Nuremberg Chronicle (1493)

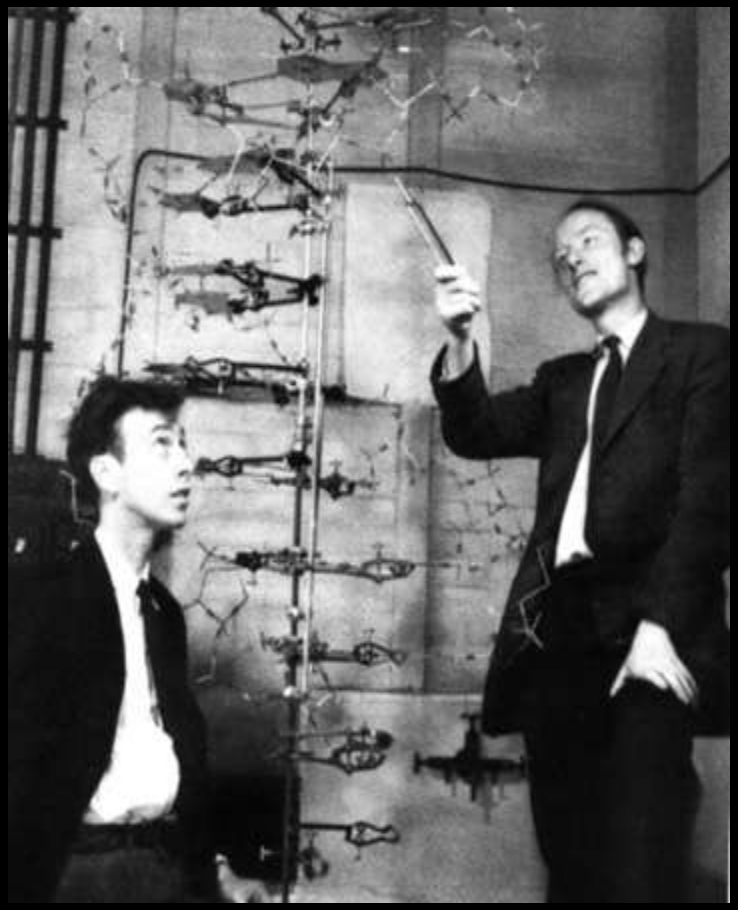


SNPs, Haplotypes and CNV Variation Across Populations



From: M. Jakobsson et al (2008) Nature 45, 999
a: percentage of allecs by geography; bb: neighbor-joining relationship tress;
c: Bayesian clustering of populations structure

From DNA Structure to Rapid, Low Cost, Massively Parallel Sequencing of Human Diploid Genomes



Nature (1953)
171, 7373
23 April

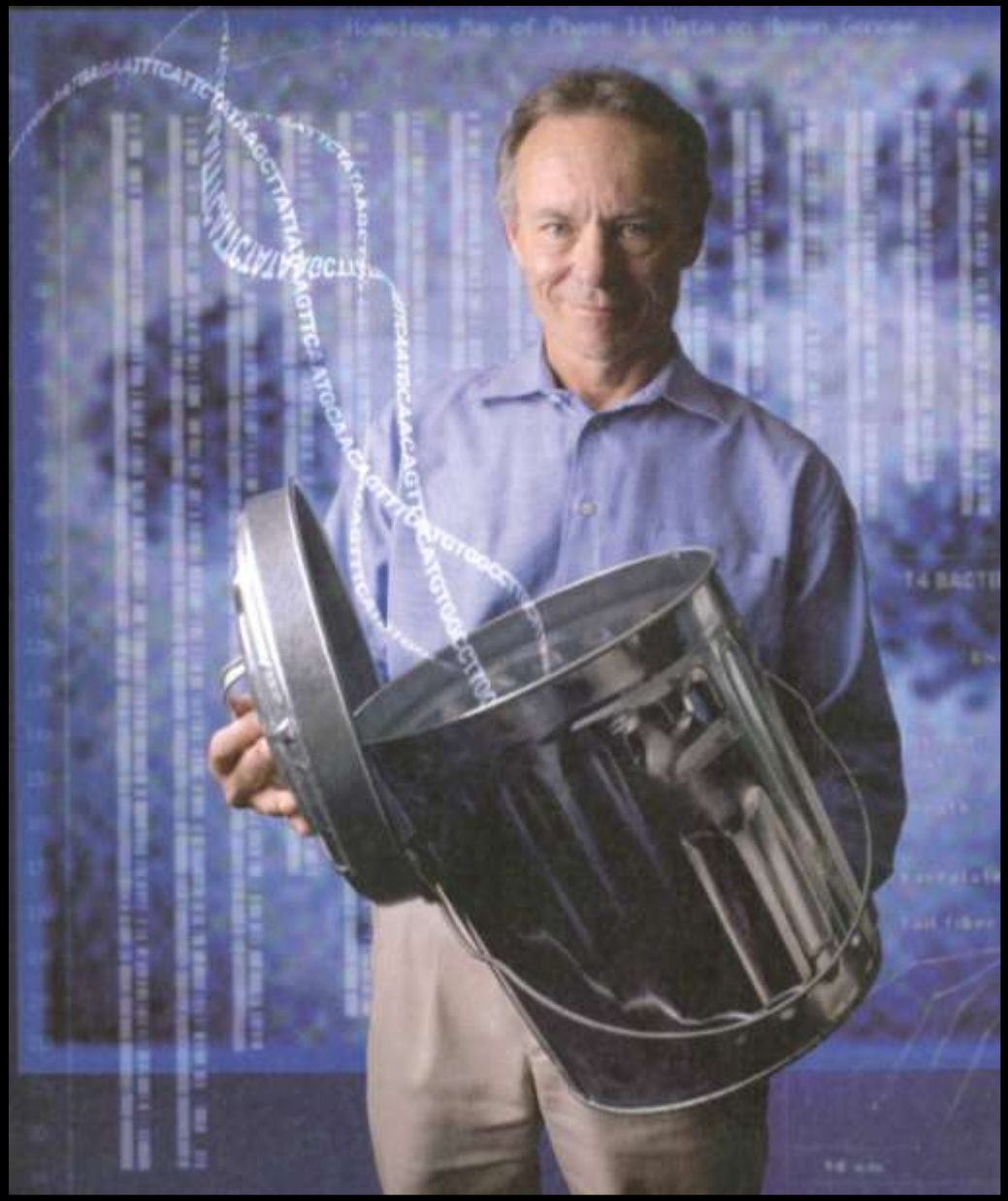
Nature (2008)
452, 872
17 April

Dr. Watson's Base Pairs

"Recognition of the thin clinical value of this sequence may cause some investors in the new sequencing methods to take pause, given the major capital investments required to commercialize these technologies have been motivated more by their perceived medical potential than by research applications."

Maynard V. Olson
Nature (2008) 452, 819

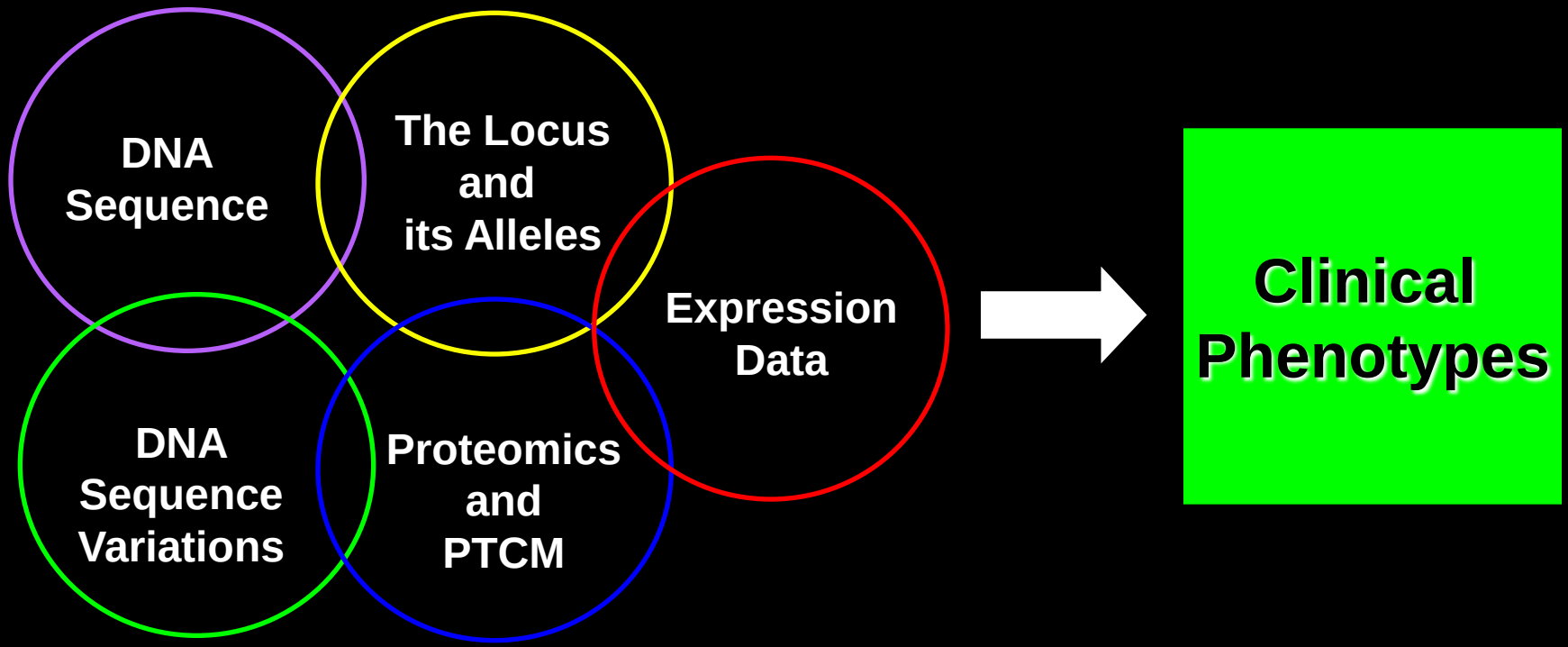
Mapping the Complexity of Genome Organization



- recognition of increasing levels of organizational and regulatory complexity
  - haplotypes
  - CNV
  - indels
  - RNA universe
  - ‘dark’ elements
  - epigenetics
  - nuclear compartmentalization and trans-expression

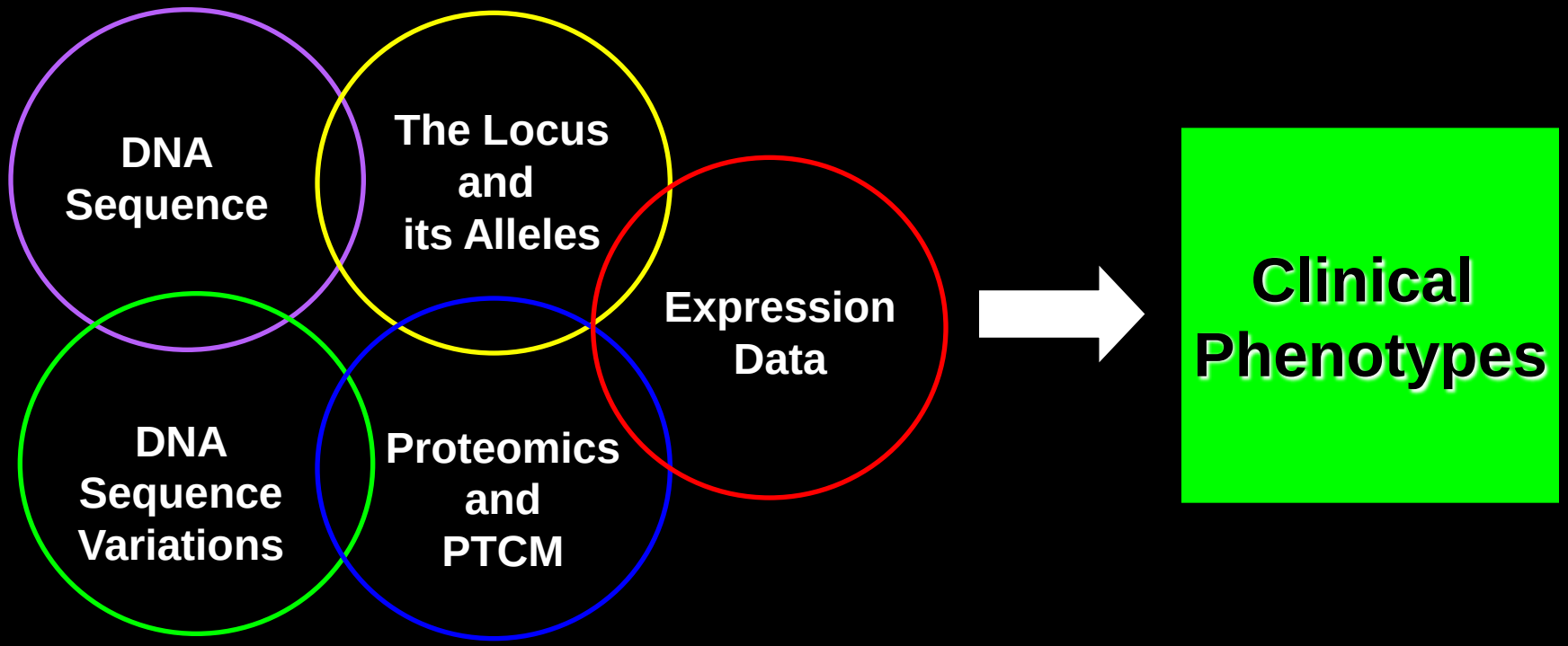
ASU BIODESIGN INSTITUTE
ARIZONA STATE UNIVERSITY

Linking Genotype and Phenotype



ASU BIODESIGN INSTITUTE
ARIZONA STATE UNIVERSITY

Linking Genotype and Phenotype



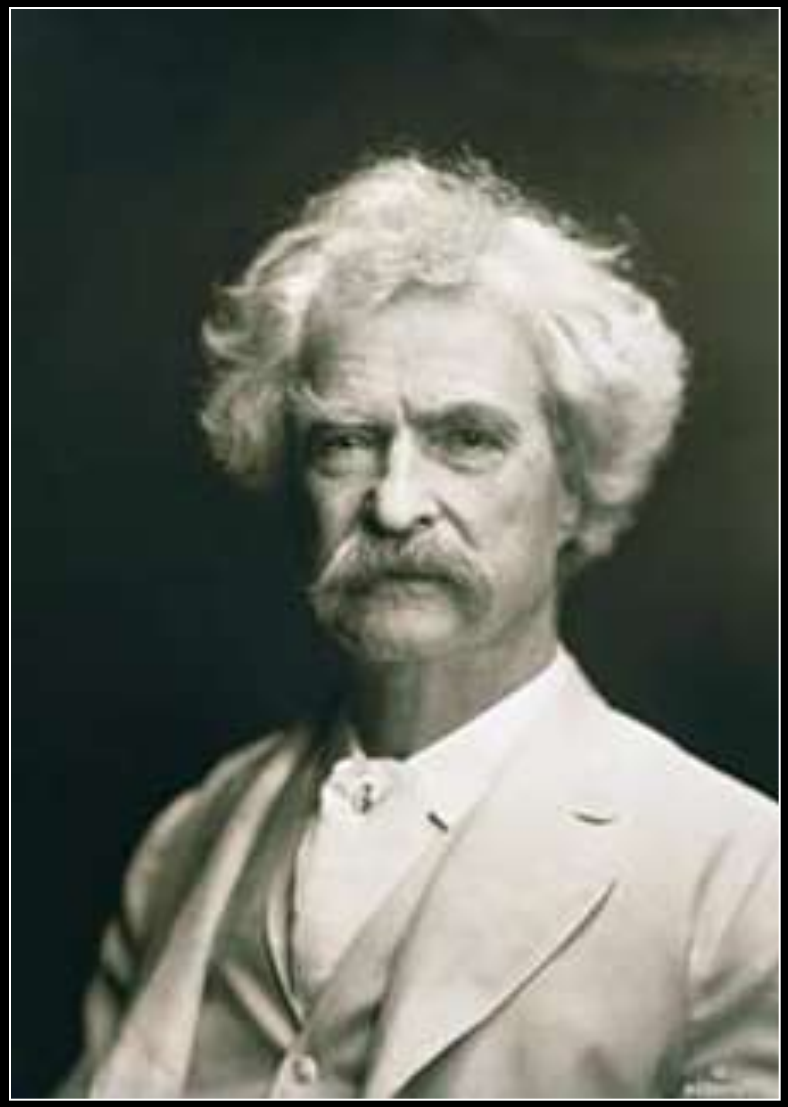
Dynamics and Plasticity of Tumor Microenvironment(s)

Proteomic-Based Biomarker Discovery

- complexity and dynamic range of plasma and other biofluids
- anticipated low relative abundance of circulating disease-specific biomarkers (> 1ng/ml)
- plasma
  - estimated 10,000 proteins
  - current methods sample less than 10% core proteome
  - detection bias towards high abundance proteins (> 1mg/ml)
- lack of amplification methods equivalent to PCR for nucleic acids
- extensive, multidimensional fractionation for analysis of complex samples
  - cost and time incompatible for clinical Dx testing
- logistical and economic ‘chokepoints’ for monoclonal antibody-based multiplex Dx

Biomarkers

- literature dominated by anecdotal studies
  - academic laboratories
  - small patient cohorts
  - limited replication and confirmatory studies
- lack of standardization
- very few biomarkers subjected to rigorous validation
  - case-control studies with sufficient statistical power
  - inadequate stringency in clinical phenotyping
- widespread lack of understanding of regulatory requirements
  - complexities imposed by multiplex tests
  - new regulatory oversight



"Supposing is good, but finding out is better."

~Mark Twain

The Analysis of Complex Biological Systems Demands A Systems-Based Approach

“Samples of Convenience”: An Intrinsic Deficit in Current Approaches to Disease Biomarker Identification and IVD Test Validation

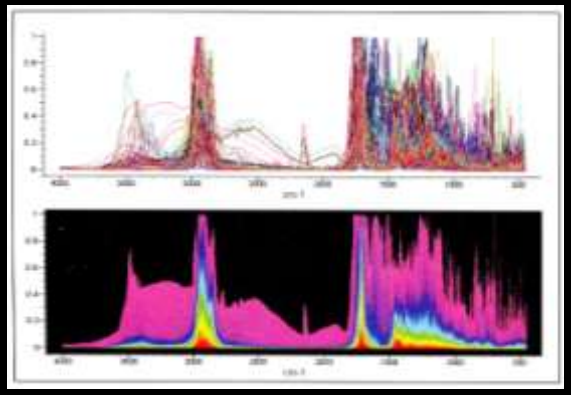
“Patients are selected based on availability of specimens, not based on representing a group for whom a therapeutic decision was made”

Dr. Richard Simon
Chief, Biometric Branch, NCI
Nature Biotechnology (2006) 24, 935

ASU BIODESIGN INSTITUTE
ARIZONA STATE UNIVERSITY

Biosignatures and Biomarkers: Chance, Bias and Exclusion of Alternate Explanations

"This field has got too much happy talk.
Biologists spend a lot of time
talking about why it should work
and not enough time figuring out 'does it work'?
Dr. David Ransohoff
UNC, Chapel Hill
Nature Biotechnology (2006) 24, 935



"During the last 30 years
biology has become a discipline
for people who want to do science
without learning mathematics"
M. Cassman et. al.
Barriers to Progress in Systems Biology
Nature (2006) 438, 1079

Sample Sizes Required to Render False Positive Results Unlikey When Testing Association Between a Genetic Variant and Cancer

Genetic relative risk†	Probability of association	Sample size for cases‡
1.25	0.001	2128
	0.0001	2580
	0.00001	3026
1.15	0.001	5789
	0.0001	7022
	0.00001	8234

from: S.G. Baker et al. (2006) BMJ 332, 1150

*Based on a two sided type I error of 0.05, a power of 0.90, and a false positive report probability of 0.05

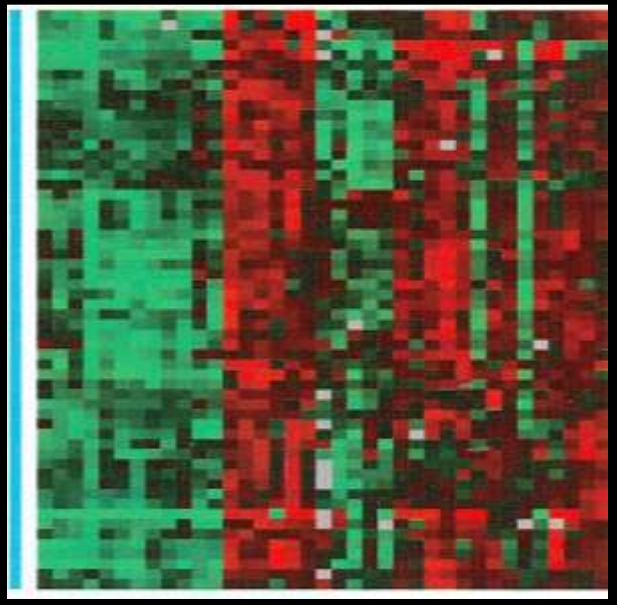
†Relative risk of cancer in people with genetic variant compared with those without.

‡An equal number of controls is also needed.

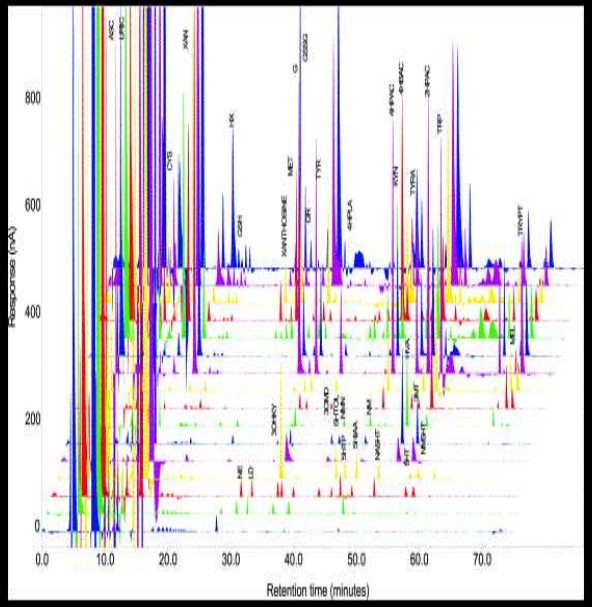
The Imperative for Rigorous Clinical Sampling Protocols in Biomarker Profiling and Validation of IVD Tests

- statistical powering
- rigorous case-control studies
  - retrospective
  - prospective (piggy back on clinical trials)
- prospectively defined endpoints
  - diagnostic marker(s)
  - Rx responsiveness and resistance markers
  - staging, stratification, progression markers
- regulatory validation of software algorithms for multiplex tests

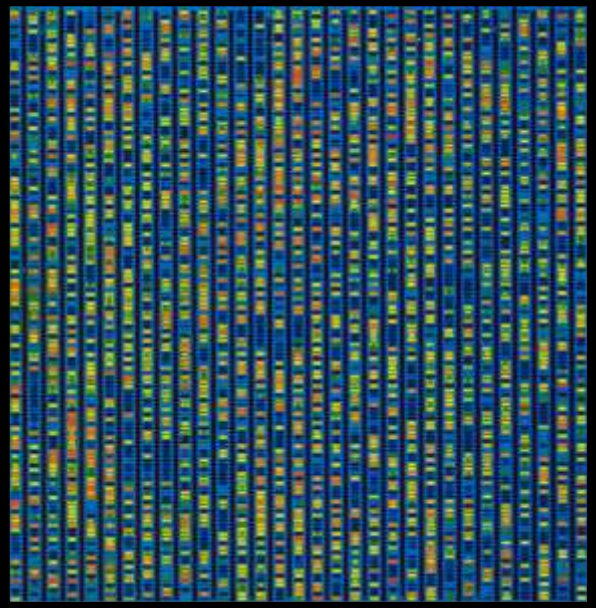
Challenges in Disease Biosignature Analysis



genomics

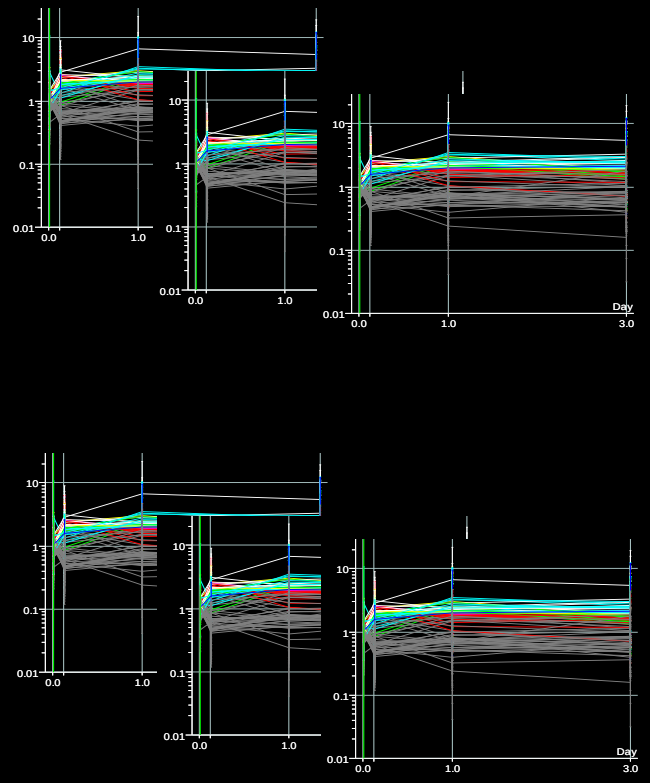


proteomics



immunosignatures

Comprehensive Biosignatures Profiling: Each Individual Serves As Their Own Control



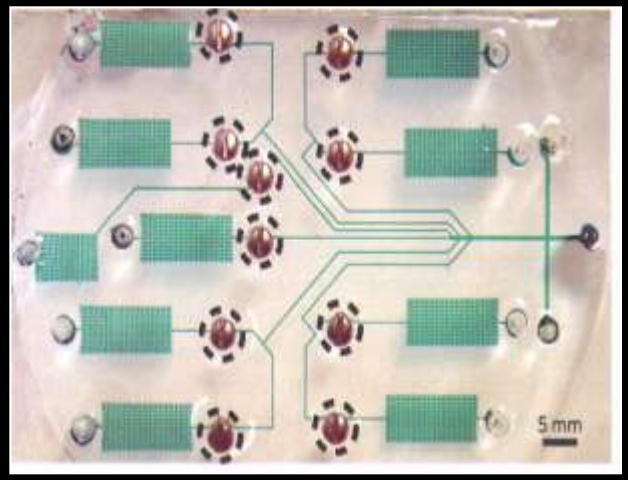
Mathematical and Computational Models of Biological Systems

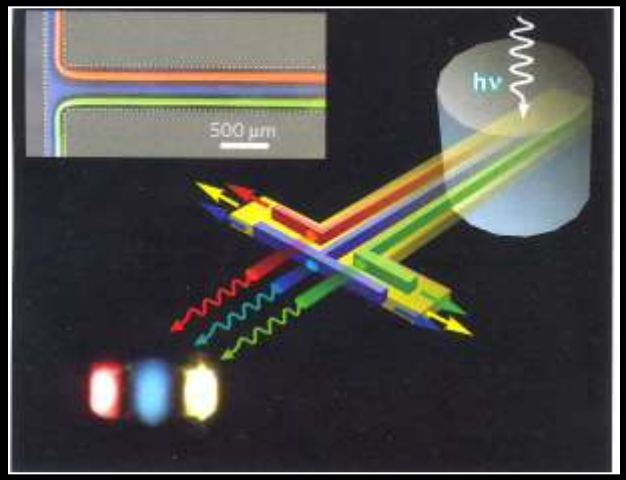
- develop precise, unambiguous and standardized representations of biological data and knowledge
- develop quantitative tools to test system dynamics in biology
- algorithmic and formal methods to address state changes, concurrency and abstraction
- identification of (bio)-logic gates in biological pathways, modules and networks
- transitioning mathematical formalisms (denotational) and computational (operational) for analysis of complex biological systems

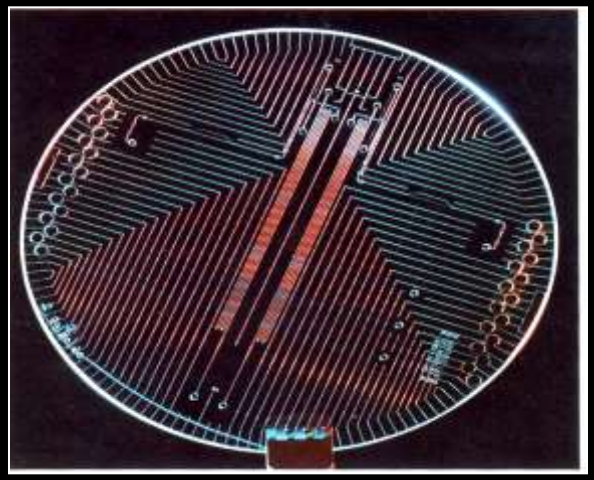
ASU BIODESIGN INSTITUTE
ARIZONA STATE UNIVERSITY

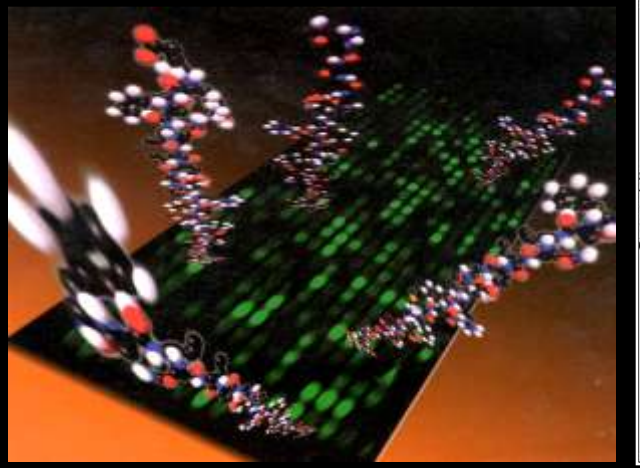
Next-Generation Multiplex Diagnostic Devices:

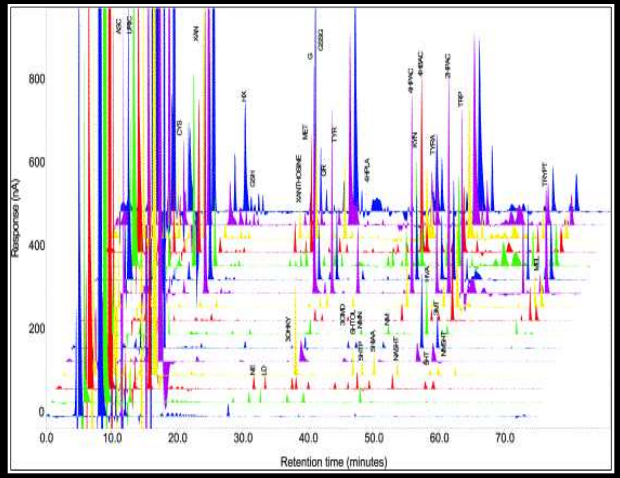
Miniaturization, Massive Parallelism, Automation and High Throughput

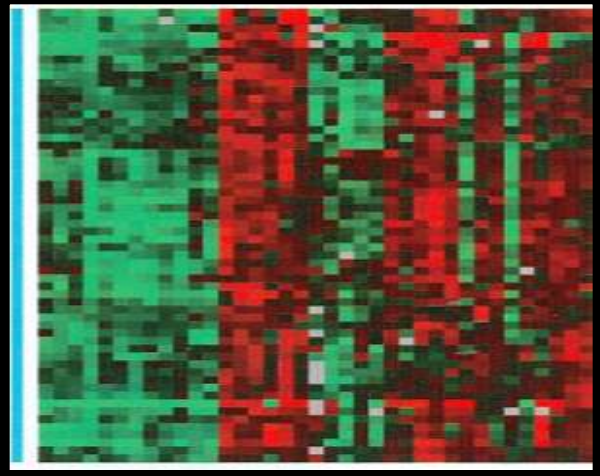








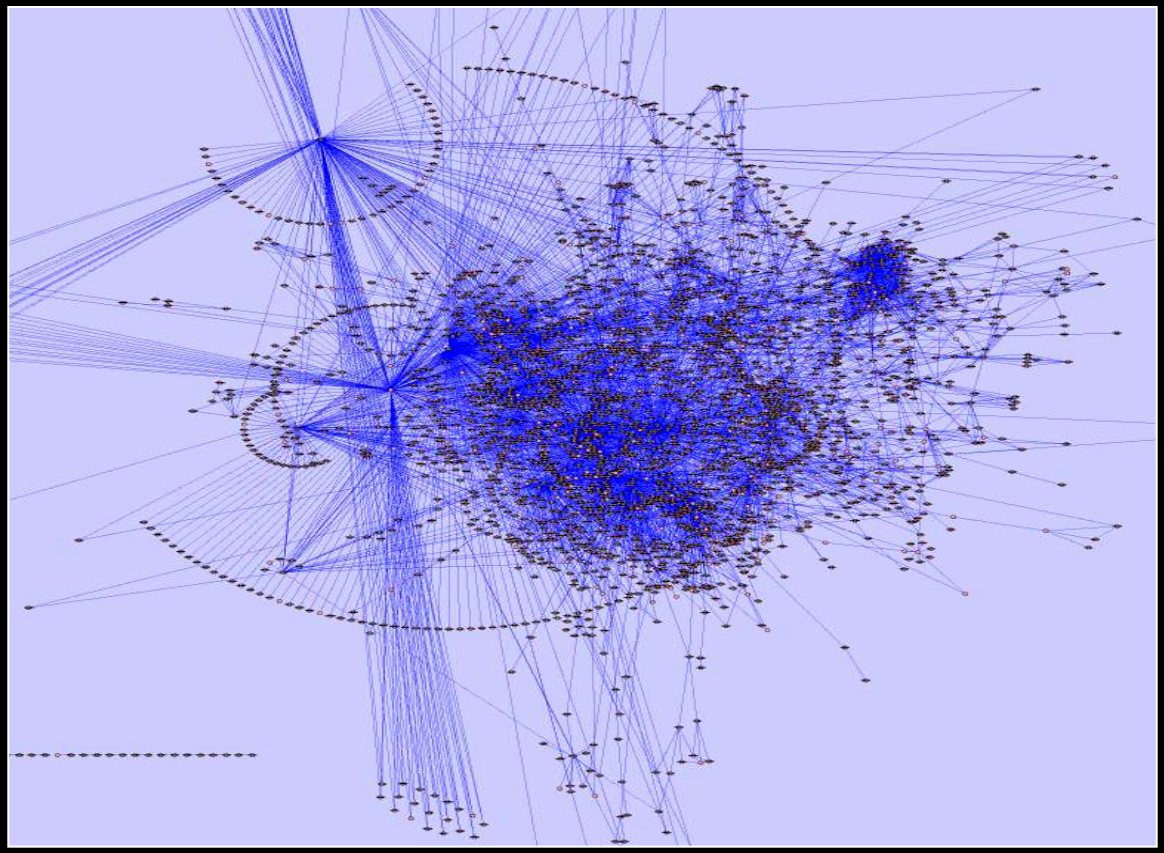




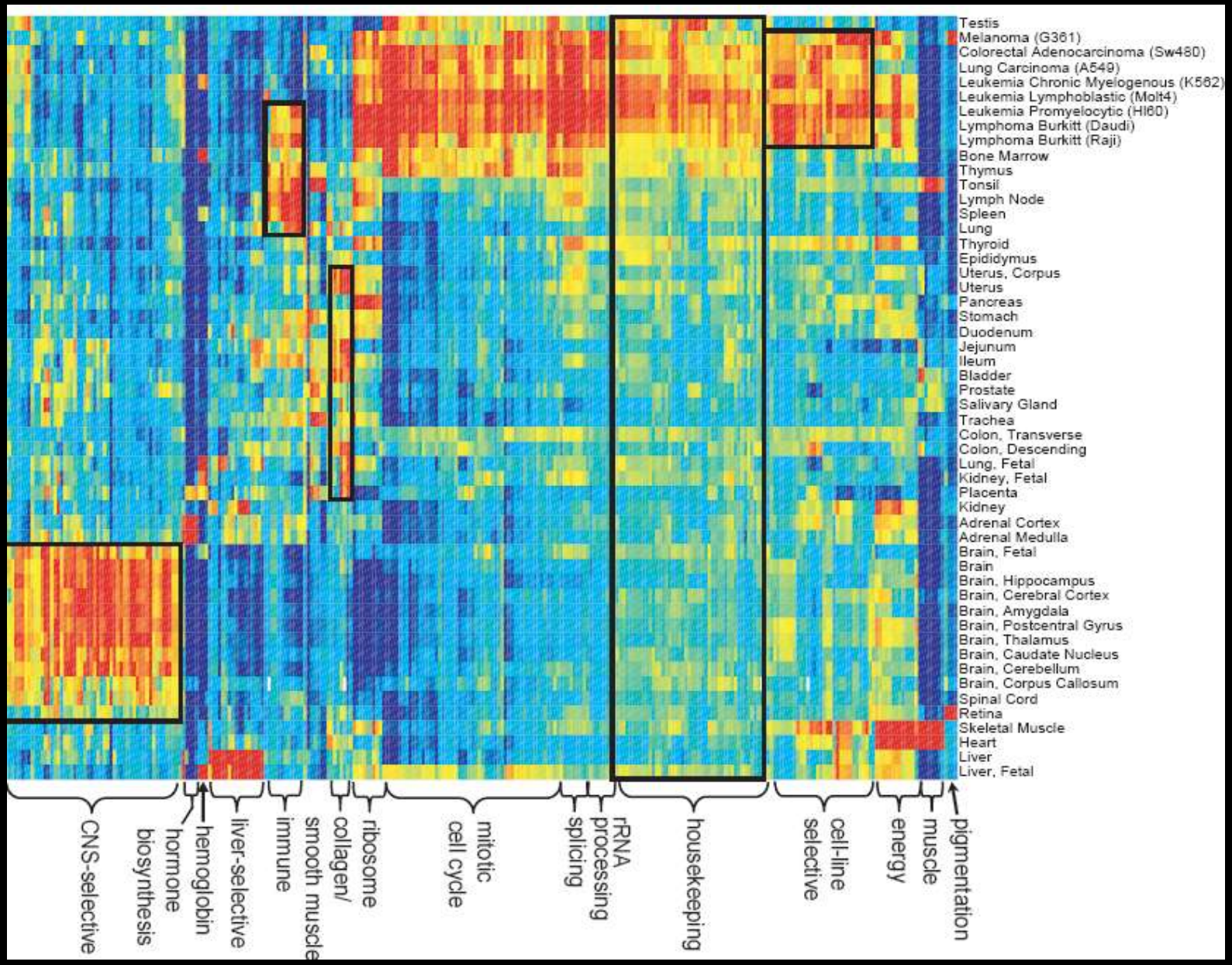
Directed Molecular Assembly

Complex Signal Detection, Deconvolution and Multivariate Analysis

Mapping the Molecular Networks of Human Diseases

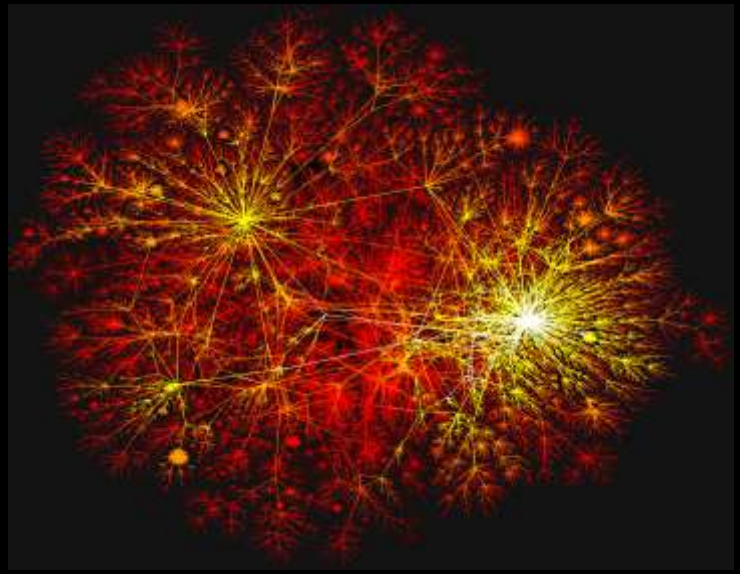


Distribution of 290 Pathway Gene Sets in 52 Human Tissues and Cell Lines

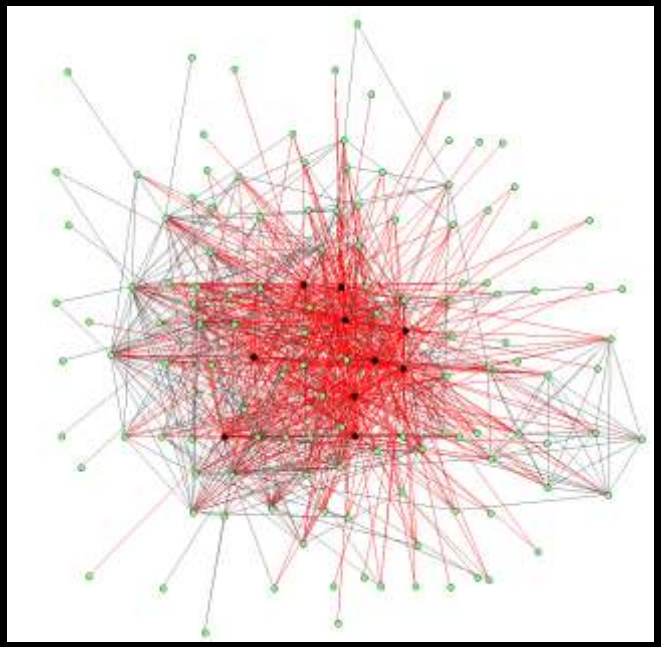


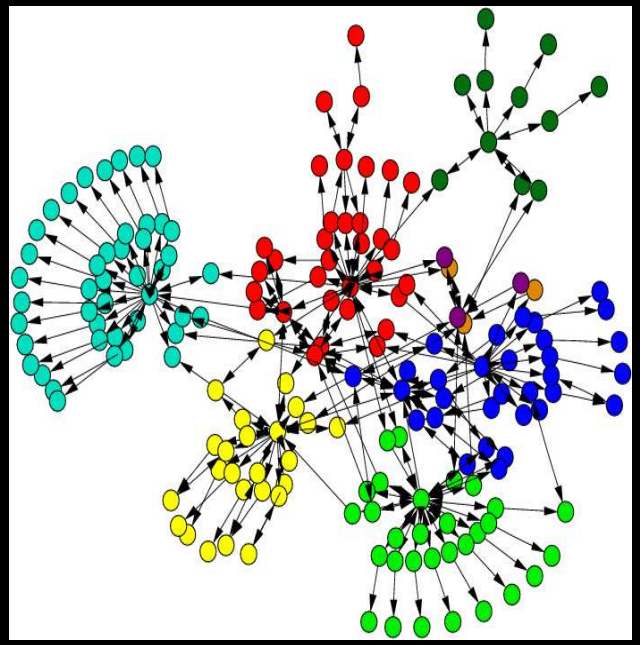
From: D. M. Levine et al. (2006) Genome Biology 7, R93

Scale-Free Network Architectures

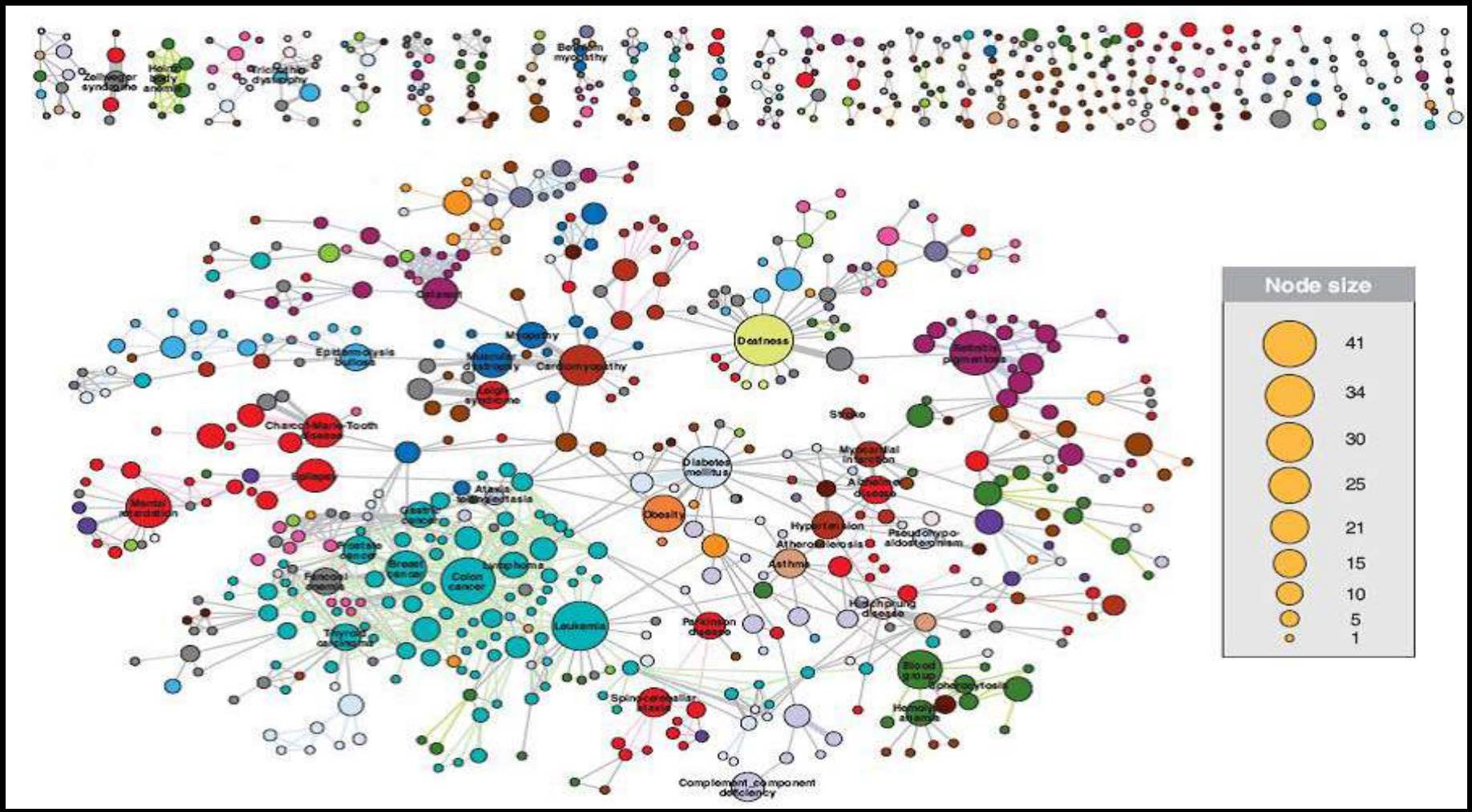


Mapping the Human Protein Interactome





The Human Disease Network
K-I Goh et al (2007) PNAS 104, 8685



- nodes define 1284 disorders and node size is proportional to # of altered genes
- link line thickness proportional to # genes shared between the linked disorders. 867/1284 disorders have at least one linkage.

Mapping the Molecular Networks of Human Diseases

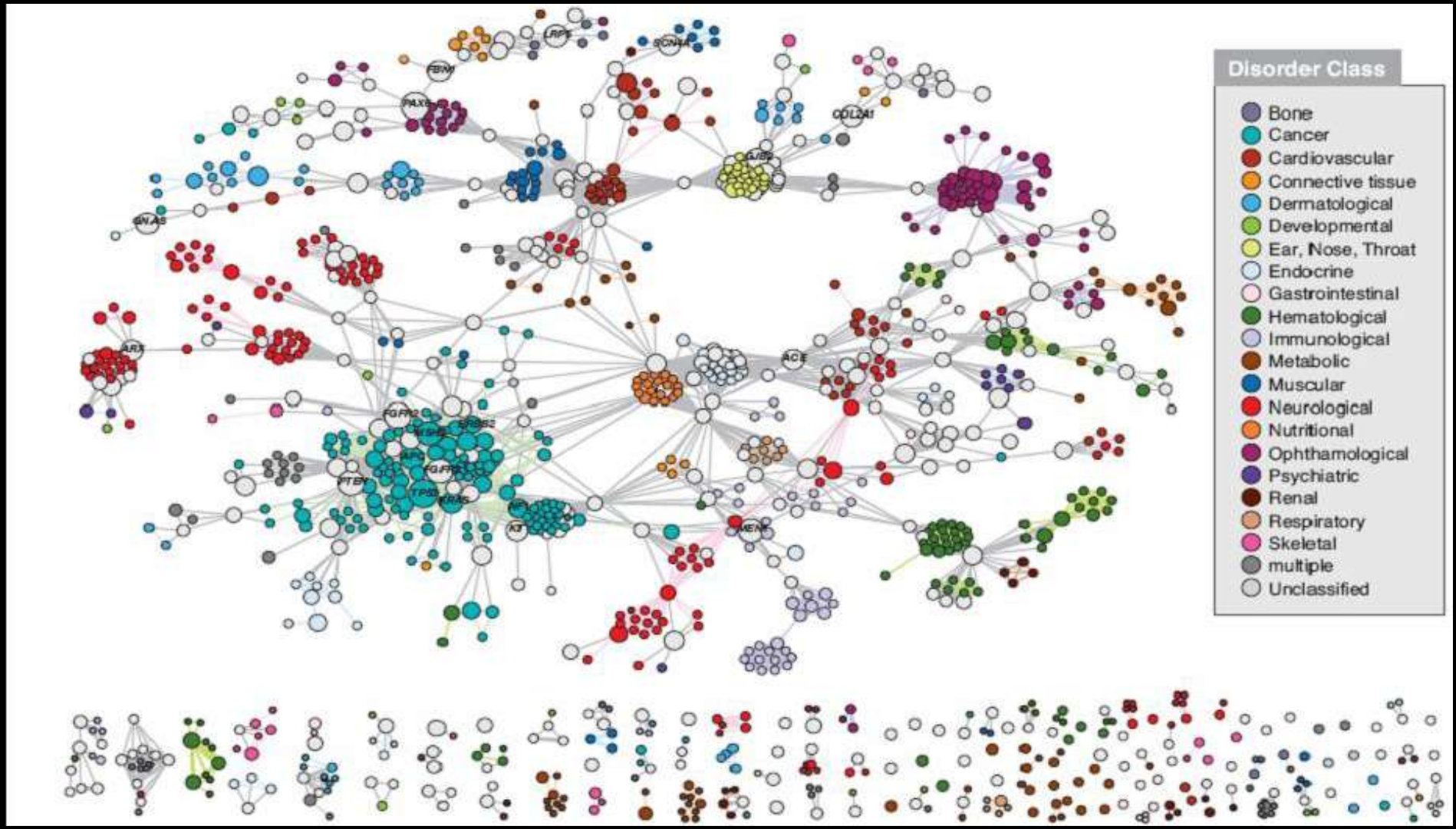
- systematic annotation of disease-associated molecular pathologies
- comprehensive inventory of gene-disease associations
- mapping the topology of gene products in biological networks
  - protein interactome analysis (hubs, intermediaries, edges)
  - characterize topological position of disease genes in specific networks and co-expression of genes
  - characterization of sites of Rx action in network topology

Impact of Molecular Heterogeneity of Disease on the Complexity of Developing Molecular Diagnostics

- extensive heterogeneity likely increases the multiplex marker spectrum needed for robust subtyping
- greater transcriptome complexity in diseases with extensive heterogeneity
  - cancer and neurological diseases exhibit highest locus heterogeneity
  - lack of correlation between expression levels and plasma proteome as potential obstacle to blood based tests?
  - potential value of new methods for capture of circulating tumor cells

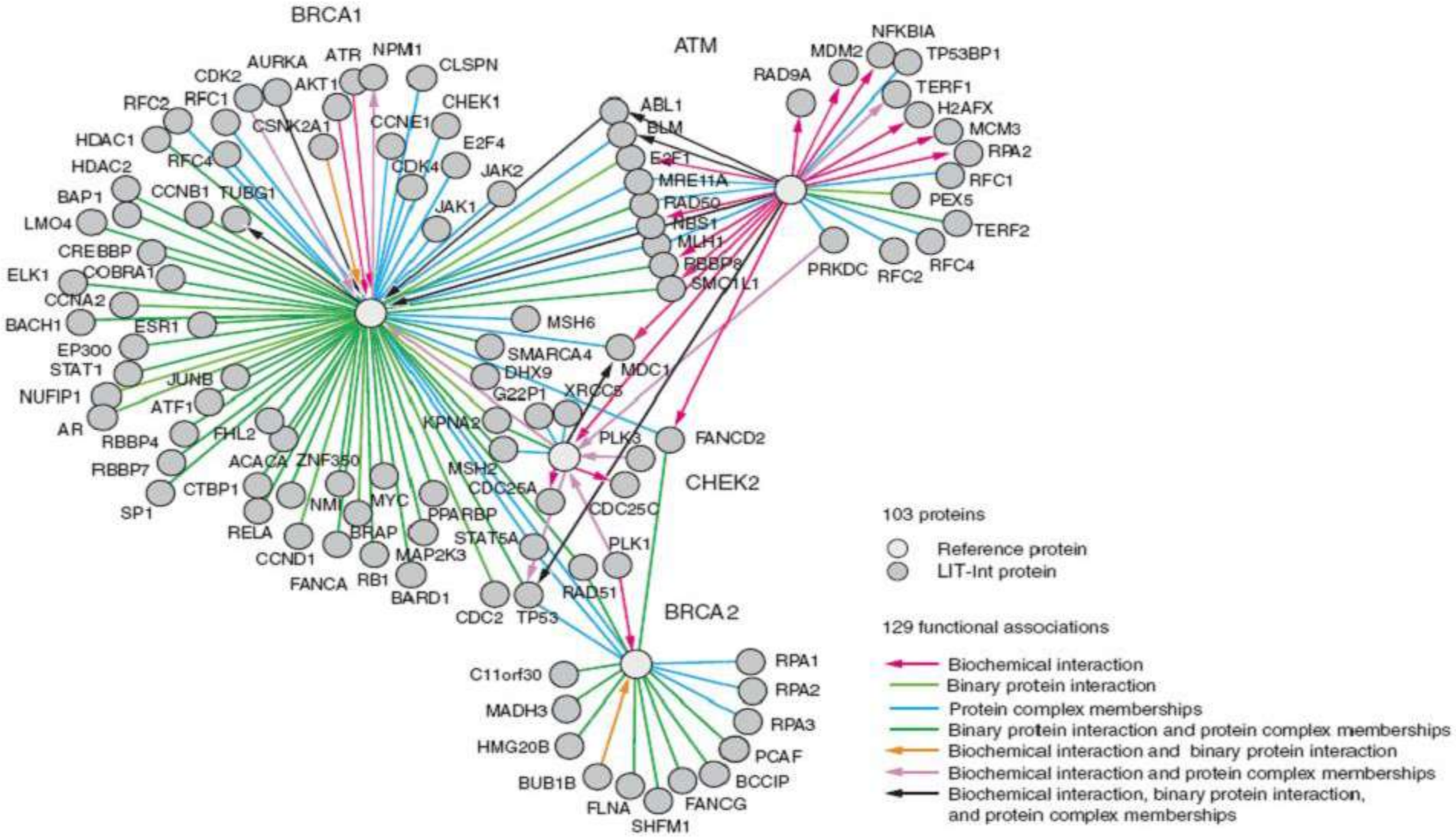
The Disease-Gene Network

K-I Goh et al (2007) PNAS 104, 8685



- nodes define specific genes and node size is proportional to # of disorders in which the indicated gene is implicated
- link lines identify genes implicated in the same disorder.

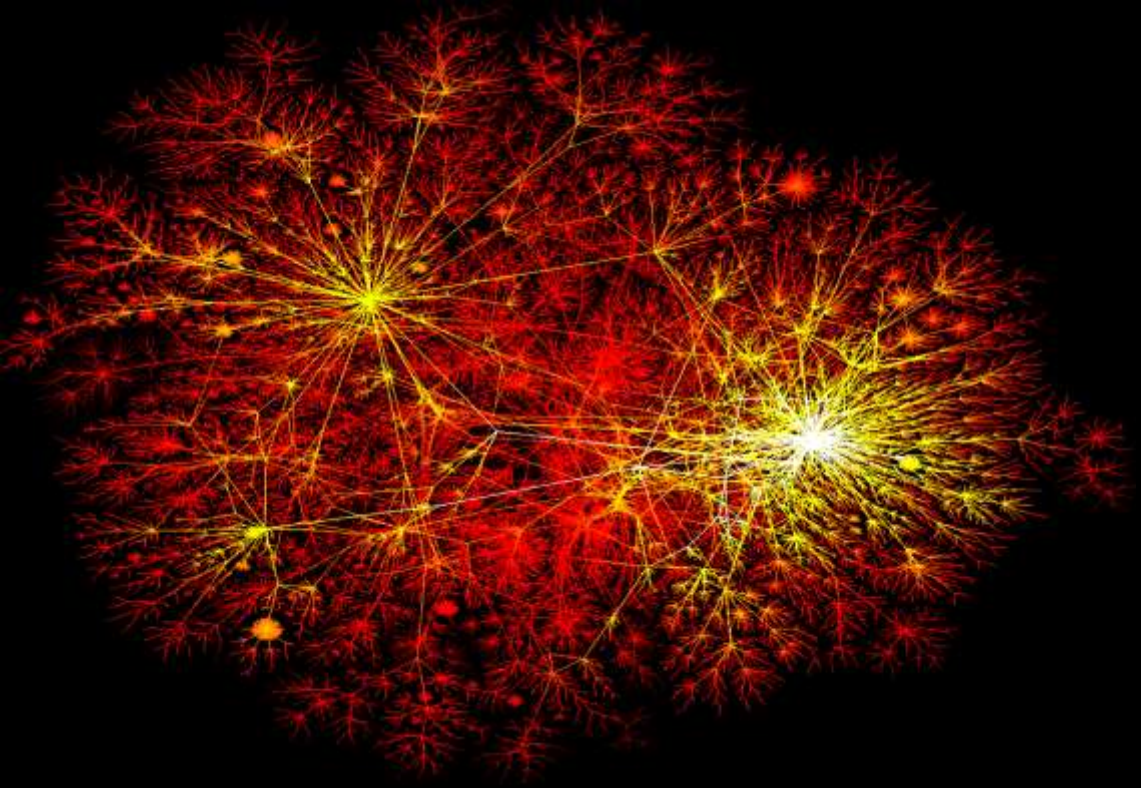
Protein Interactome Network Analysis of Two High Penetrance (BRCA 1/2) and Two Low Penetrance (ATM, CHEK) Genes Implicated in Human Breast Cancer



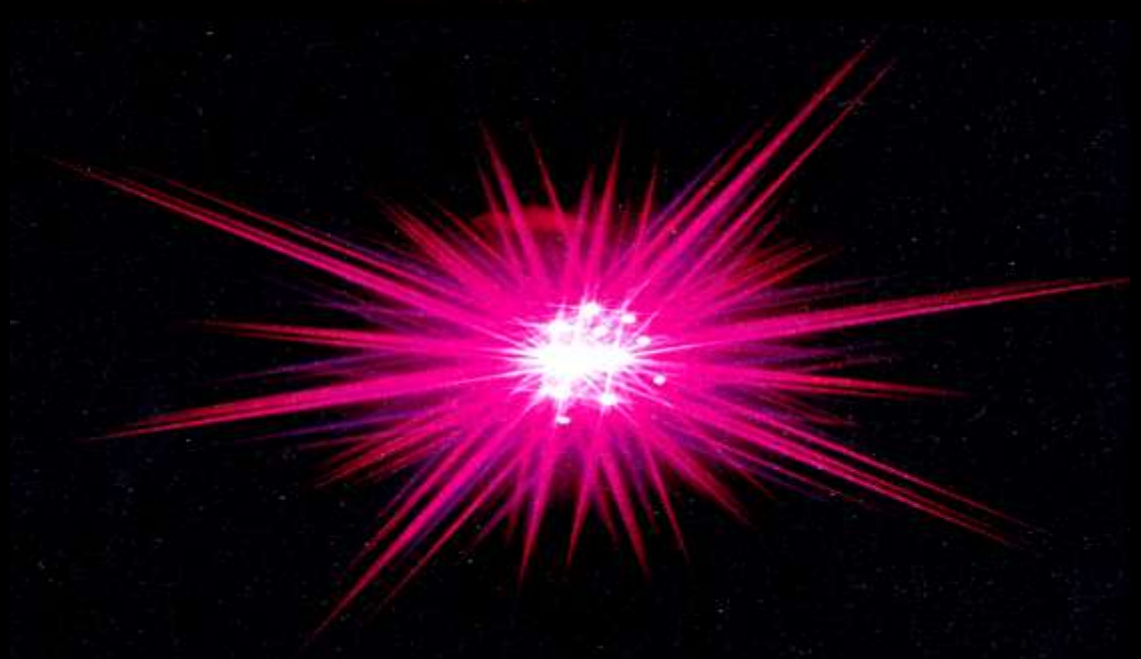
From: M. A. Pujana et al (2007) Nature Genetics 39, 1378

Human Disease Networks and Disease-Gene Networks
K.I. Goh et al (2007) PNAS 104, 8685

- 1377 of 1777 disease-associated genes involved in more than one disease
- most diseases characterized by alterations in relatively low number of genes
- cancer and neurological diseases exhibit changes in large number of genes and high levels of connectivity between the implicated genes
- metabolic, skeletal and multiple disorders exhibit low genetic heterogeneity and low disease-gene 'connectivity'



Deconvolution of Signaling Networks in Disease



Identification of ‘Fragile’ Nodes/Pathways for Targeted Rx

Network (Systems) Pharmacology

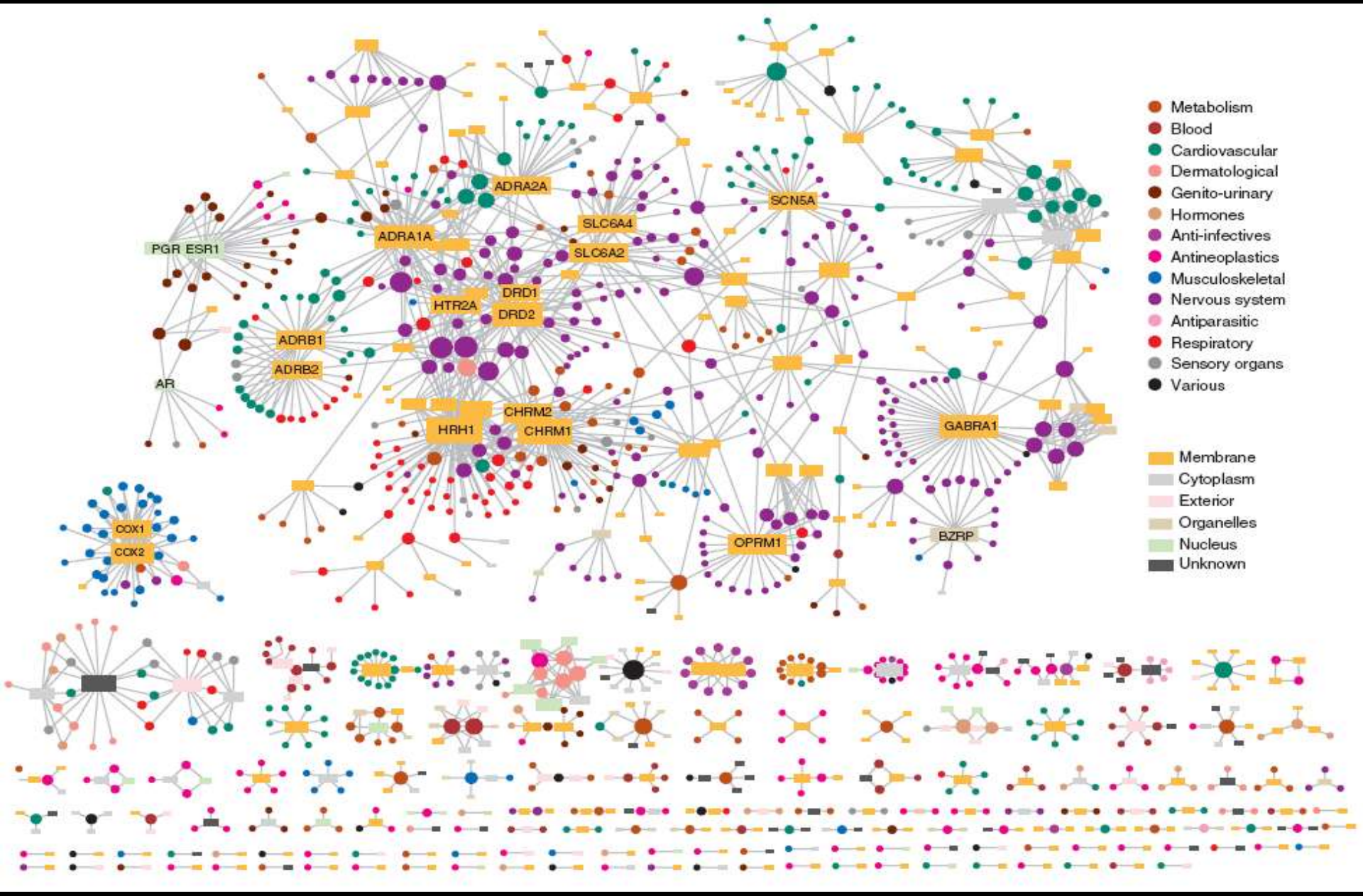
drug-target network

- map Rx action on different targets

target-protein network

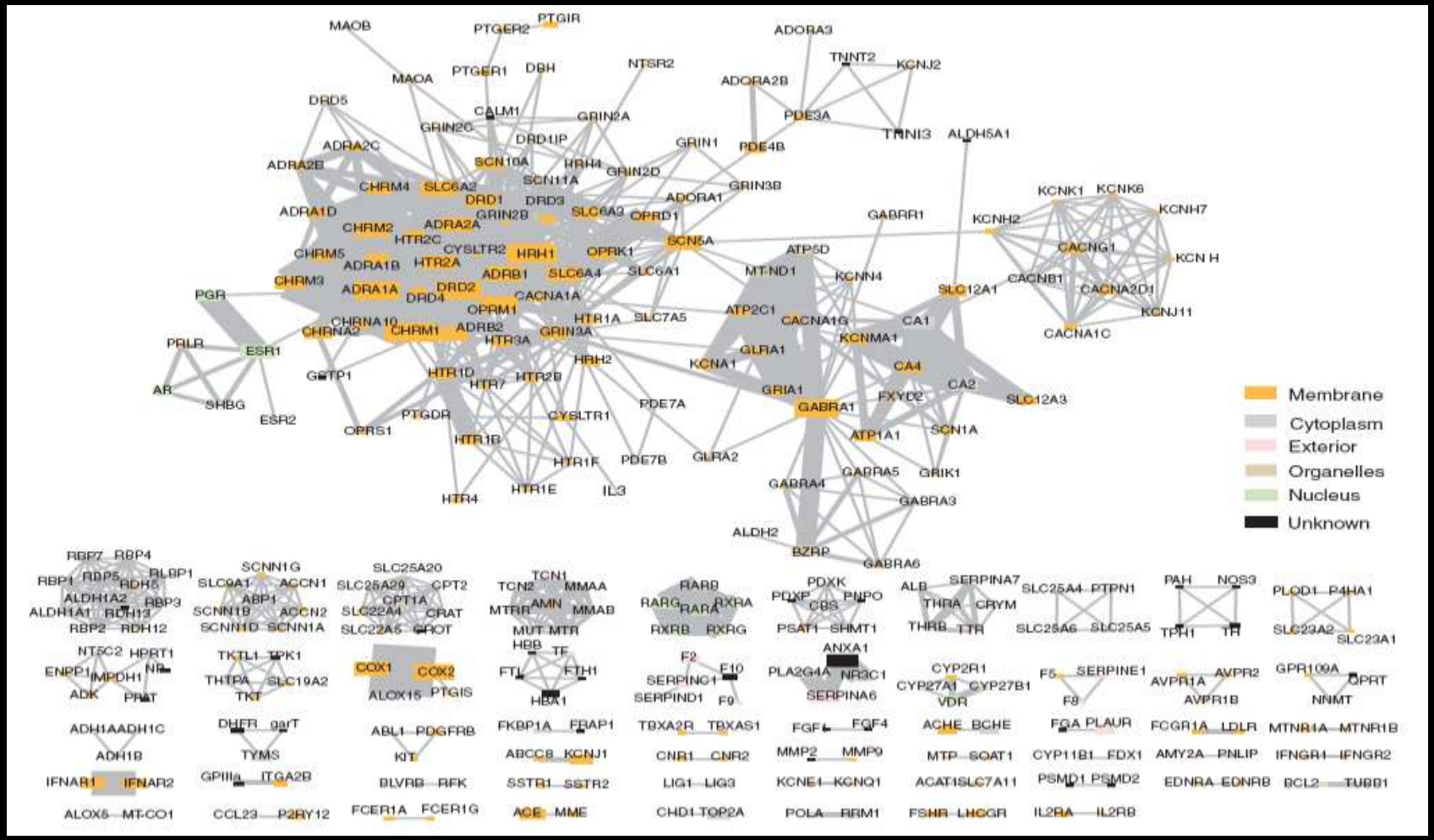
- identify two or more proteins targeted by at least one common Rx

Drug-Target Networks for FDA Approved Rx:



M. A. Yildirim et al. (2007) Nature Biotechnology 25, 1119

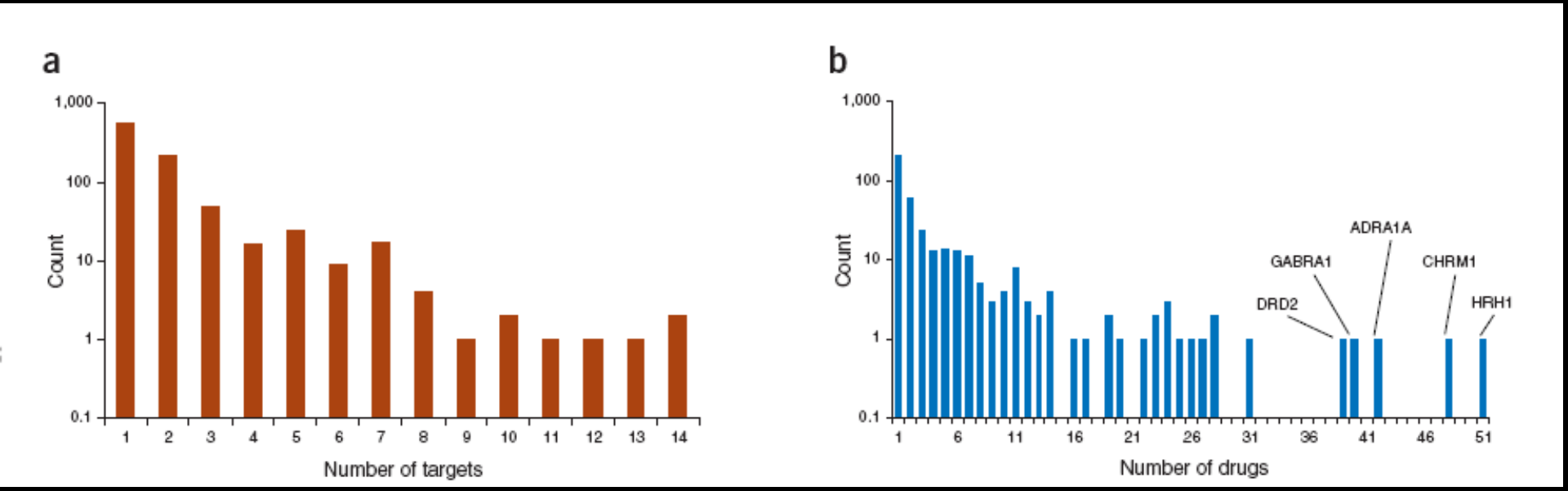
Target-Protein (TP) Network for FDA Approved Drugs: Connectivity of Proteins Targeted by the Same Drug(s)



Thickness of connector proportional to # drugs acting on the same target

From: M.A. Yildirim et al (2007) Nature Biotechnology. 25,1119

Drug-Target Interaction Map for 1178 FDA Approved Drugs (2006)

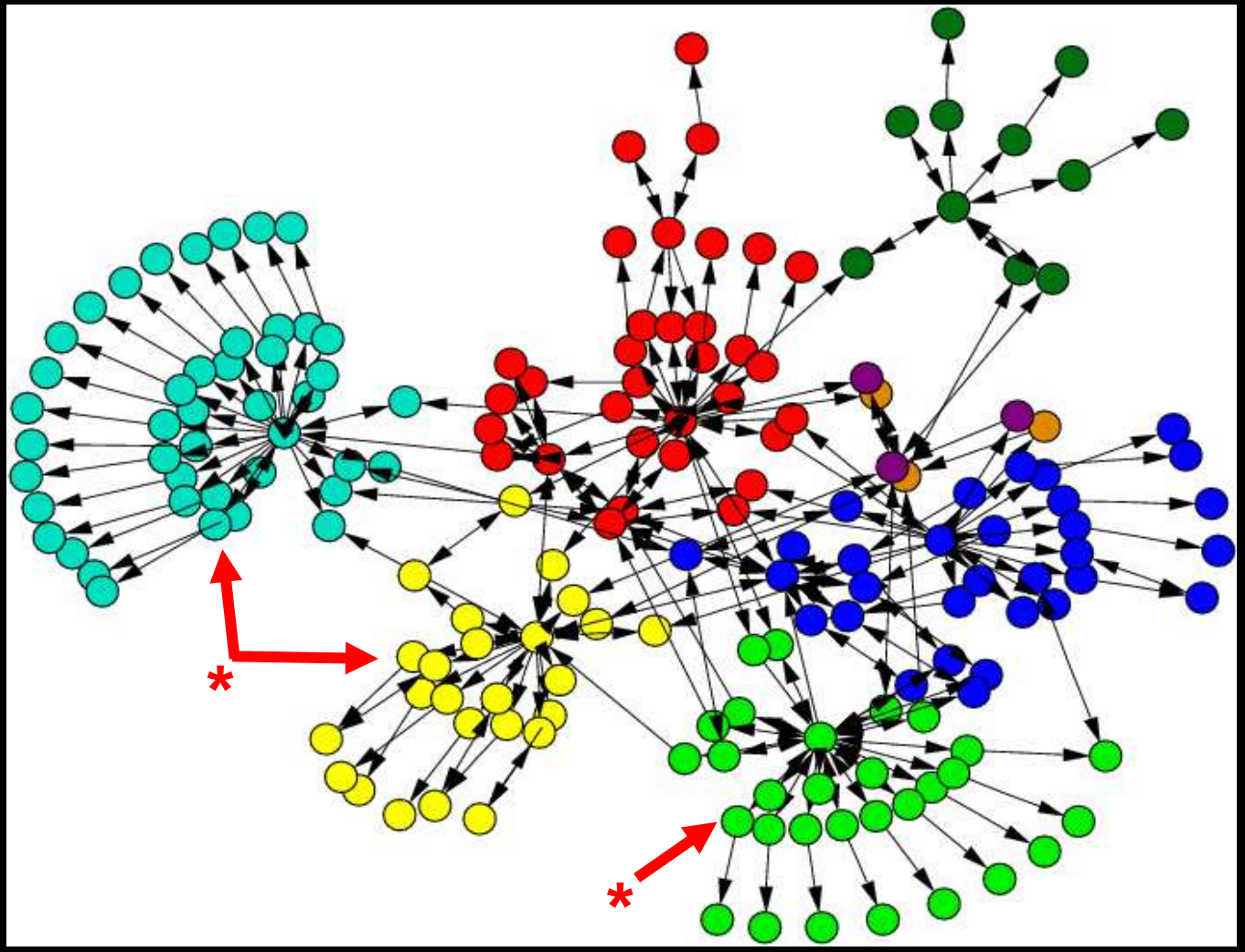


From: M.A. Yildirim et al (2007) Nature Biotechnology. 25,1119

a = distribution of drugs with respect to # targets

b = # of drugs that target specific proteins

Emerging Features of Systems-Level Determinants of Molecular Targets for FDA-Approved Drugs



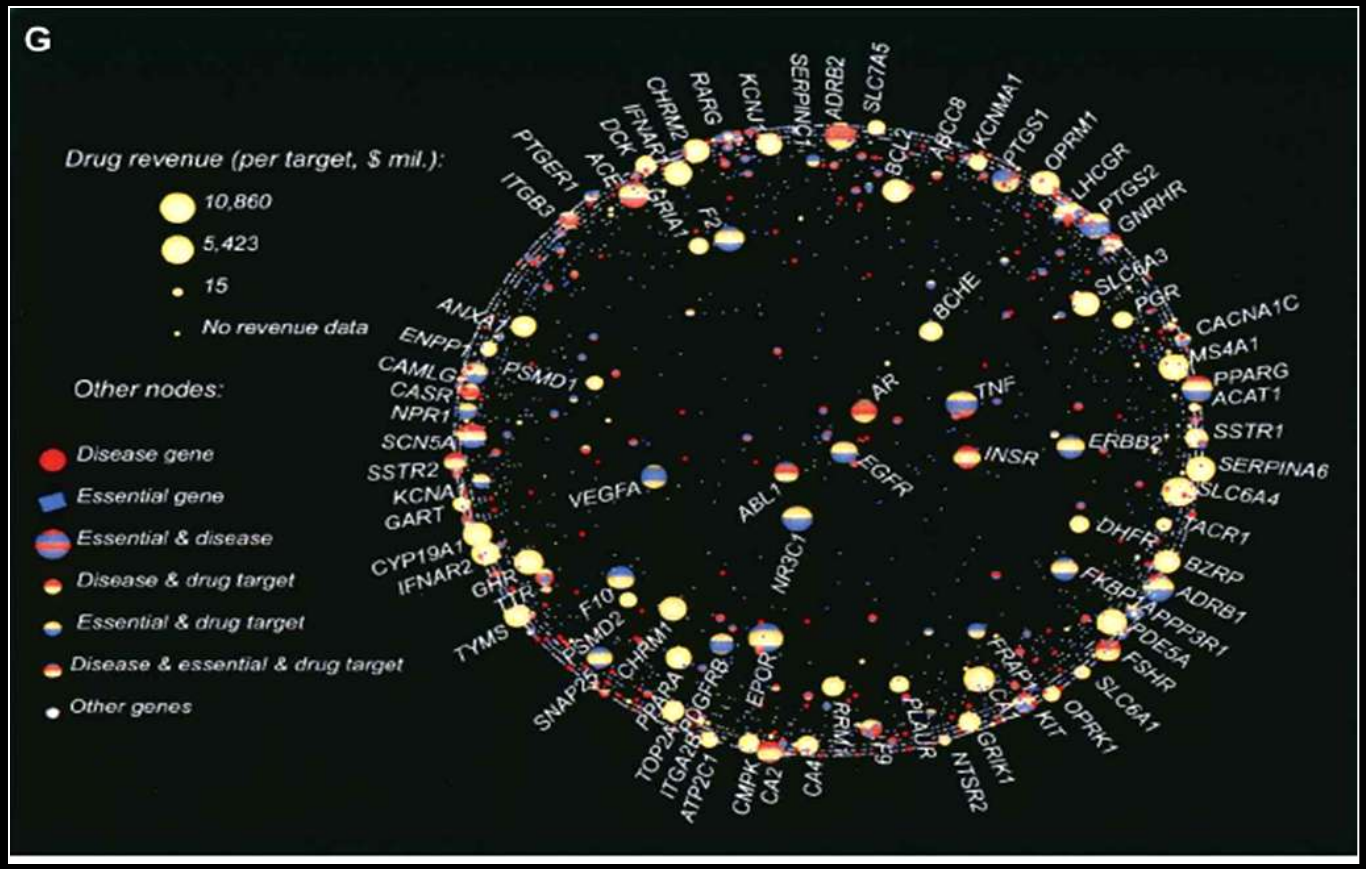
Emerging Features of Systems-Level Determinants of Molecular Targets for FDA-Approved Drugs

- Rx targets occupy narrow niche in this distribution
  - connectivity greater than average node in the network

    but

    small compared to maximum connectivity exhibited by hubs
  - less nonsynonymously polymorphic than human genes on average across the network
  - lower Shannon entropy expression scores than average (i.e. higher tissue specificity)

Plot of Rx Target Connectivities Superimposed with Rx Revenue Data



From: L. Yao and A. Rzhetsky (2008) Genome Res 18, 206
Closer a target is to the center the higher its network connectivity. Color of node conveys functional status and size reflects revenue.

Systems Biology and the Design of "Promiscuous" Multi-Target Therapeutics

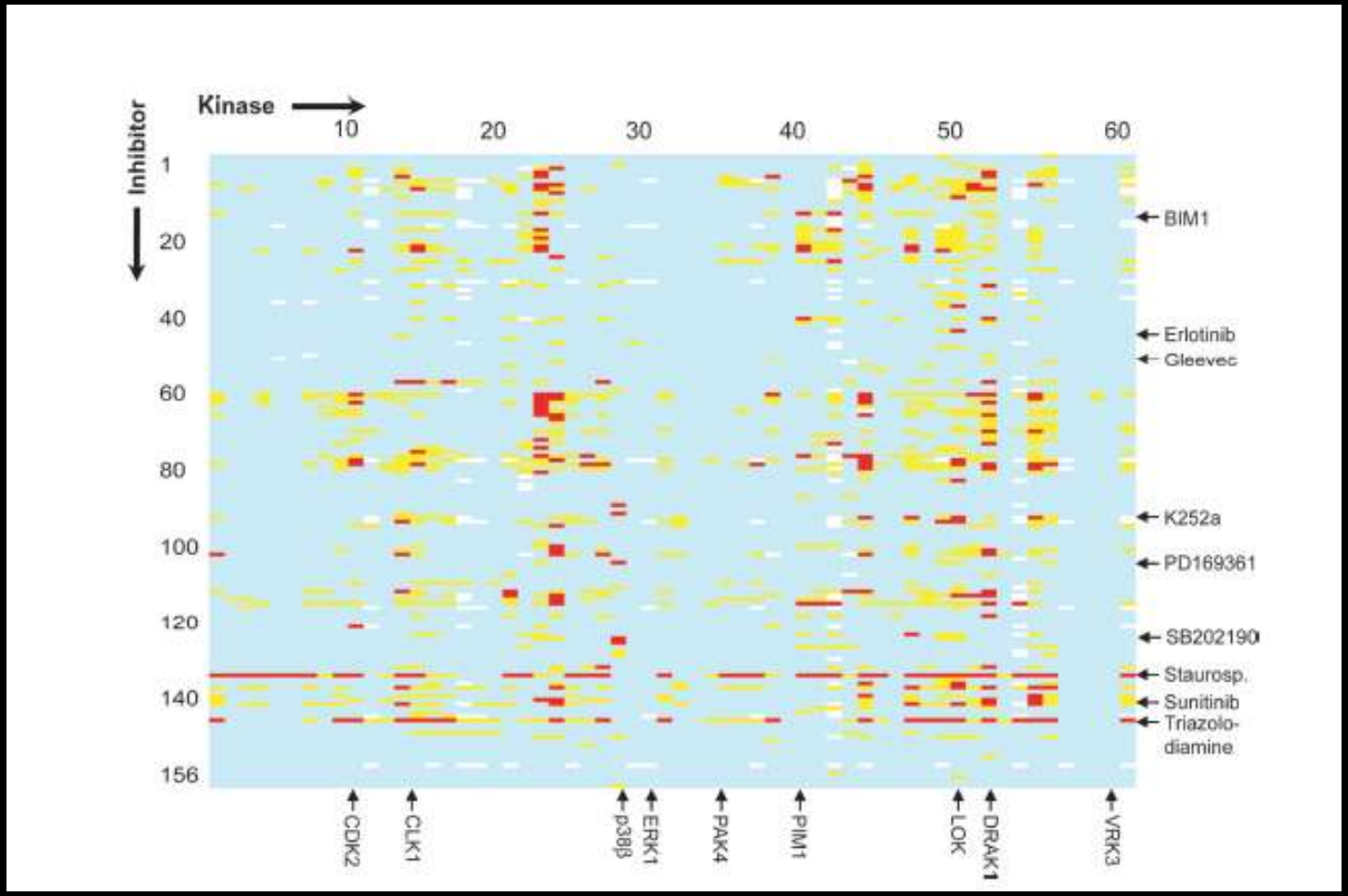
- impact of reductionism of molecular biology/genomics on Rx discovery
  - "one target : one drug" strategy
  - rejection of promiscuous candidates in HTS screening
- robustness and redundancy in biological networks likely requires effective Rx interact with several proteins
  - designed multi-ligand (DLM) agents
- identification of relevant "multi-targets" for Rx action
  - 2P interactions, siRNAs, multi-knockouts

ASU PROPRIETARY

Network (Systems) Pharmacology

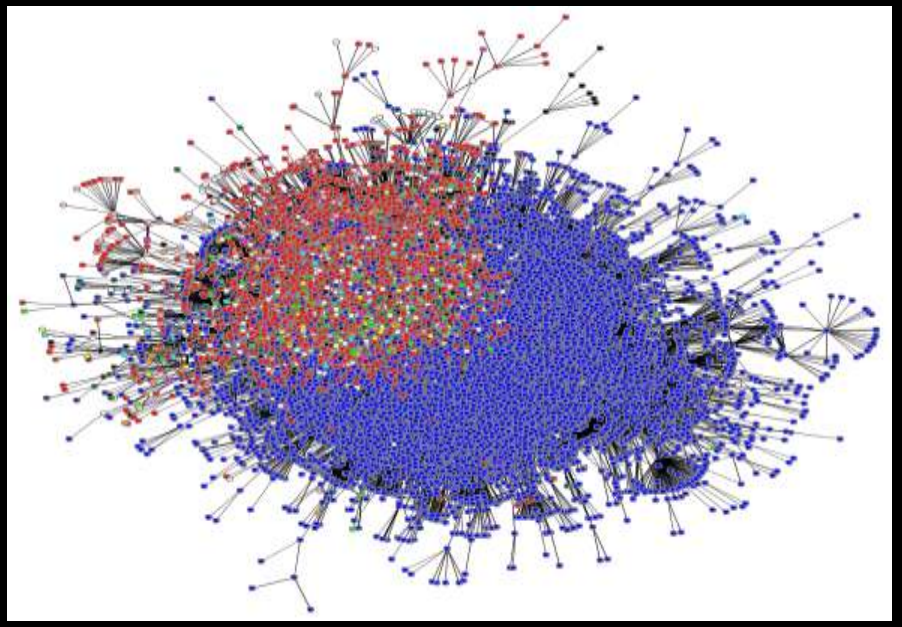
- analysis of Rx safety/efficacy in context of biological pathways/networks
- shift from one drug:one target strategy to recognition of benefits of polypharmacology and target promiscuity in Rx action
- exquisitely selective Rx may exhibit lower than desired efficacy
- Rx acting on two or more targets in disease pathway may be more efficacious
- parallel insight that same Rx target may be implicated in multiple diseases
  - new opportunities for Rx combinations?
  - new combination Rx approaches with lower doses of each Rx?

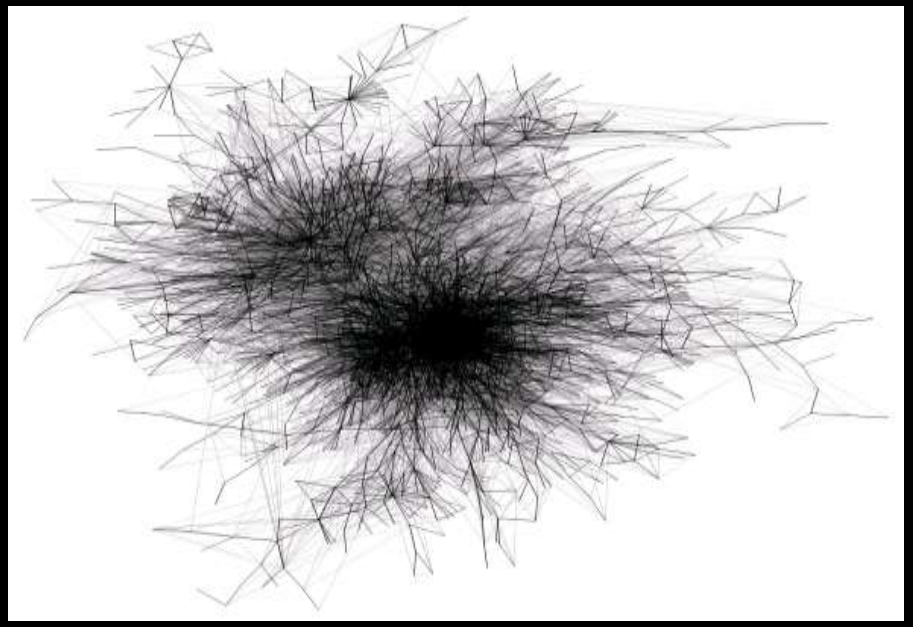
Interaction Map of Kinase Inhibitors with Ser/Thr Kinases Using Thermal Shift Assay

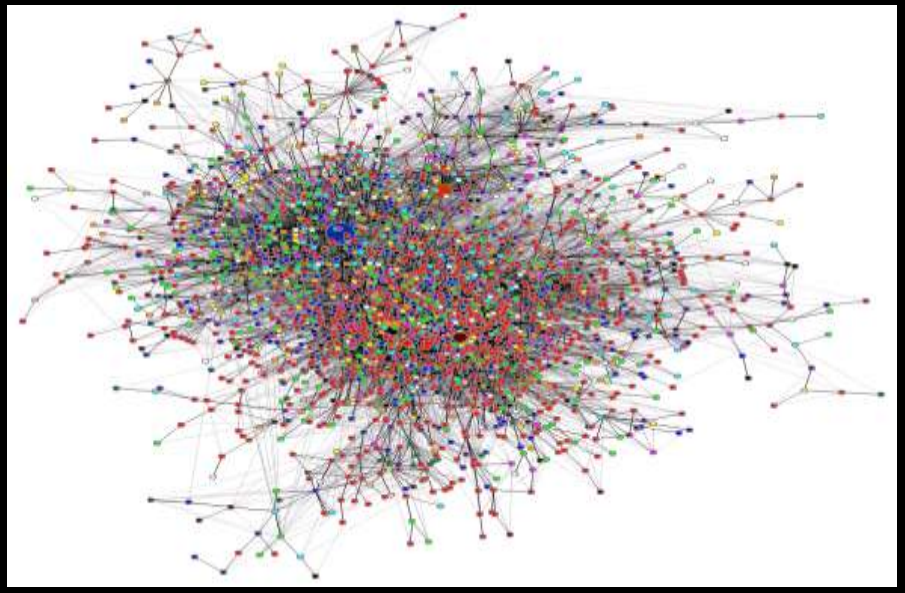


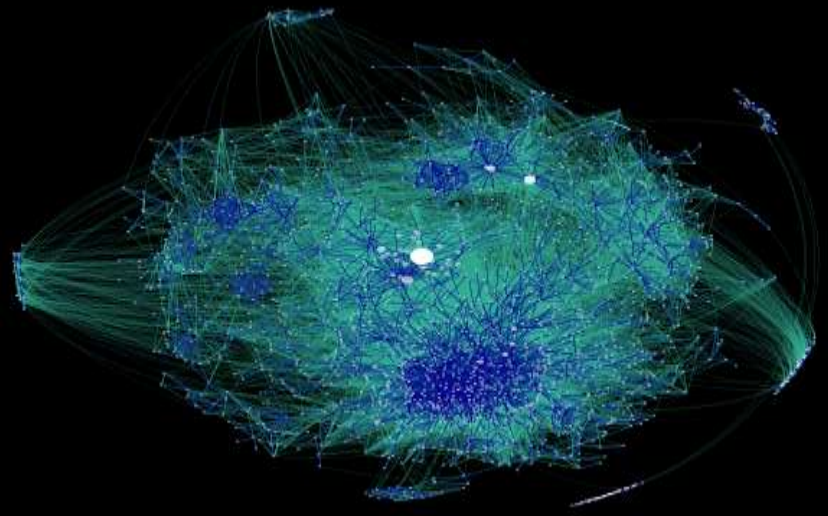
Yellow = Tm > 4° C (Kd < 1μM) red = >8 ° C (Kd< 100 nM)
From: O. Fedorov et al (2007) PNAS 104, 20523

Deconvolution of Complex Networks

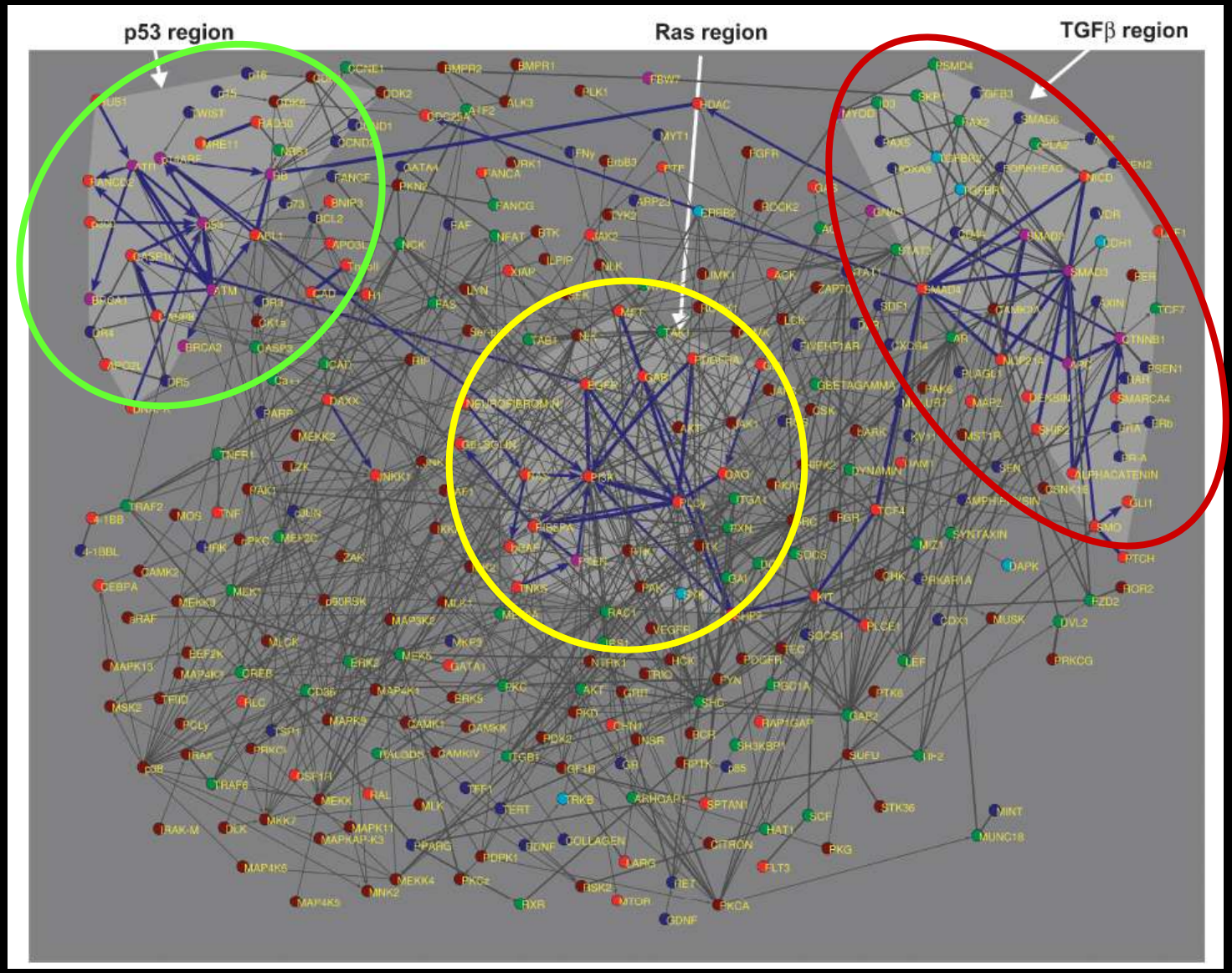








Human Oncogene-Signaling Map



Adapted from: Q. Cui et al (2007) Mol. Sys. Biol. 3, 152
326 nodes, 92 links and 12 topological ‘blocks’

Signaling Pathways in Human Cancers

- comprehensive inventory of network topology of cancer-associated genetic (mutation) and epigenetic alterations
- mutated genes are enriched in positive-signaling (activation) regulatory loops
- methylated genes are enriched in negative-signaling (suppressor) loops
- activating pathway changes occur in signaling hubs rather than nodes with structural/scaffold functions

Signaling Pathways in Human Cancers

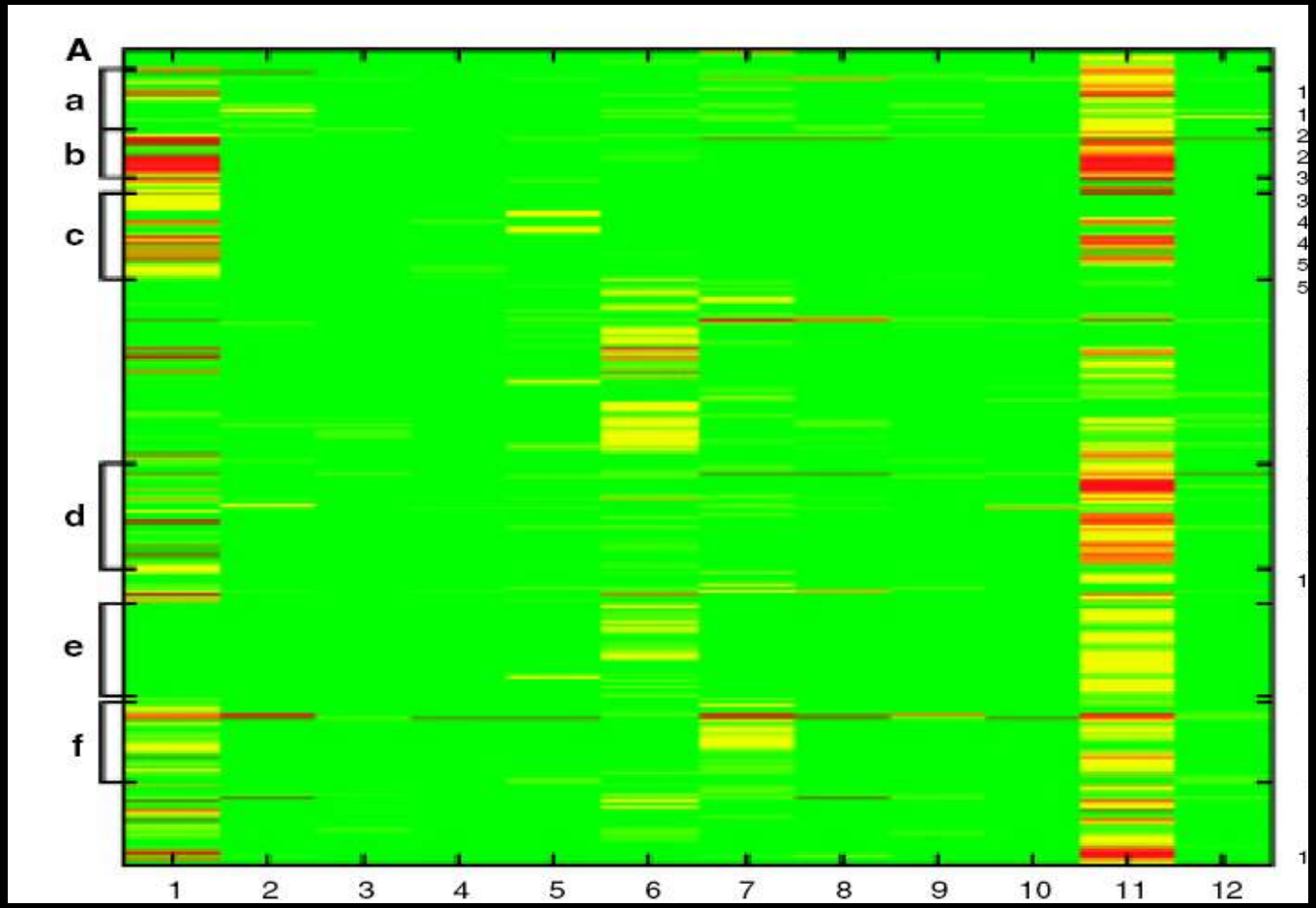
- topology of oncogene and suppressor gene networks
- highly connected regions (blocks)
- emergent patterns of functional collaborations between genes in subsets of different blocks in tissue-dependent fashion

Heatmaps of Gene Mutation Distributions in 12 Oncogene-Signaling Blocks

breast
CNS
blood

lung

pancreas
skin

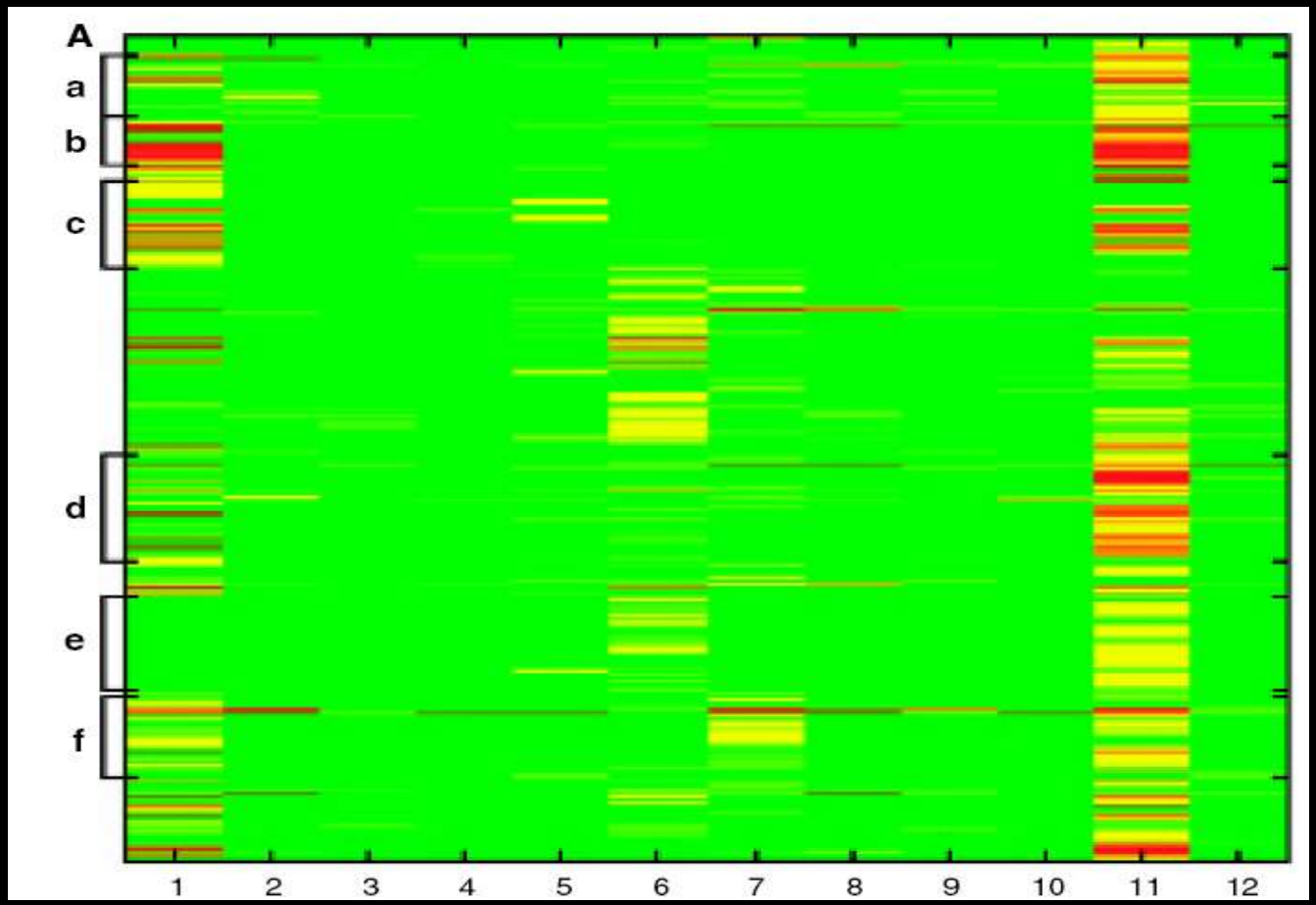


From: Q. Cui et al (2007) Mol. Sys. Biol. 3, 152

Heatmaps of Gene Mutation Distributions in 12 Oncogene-Signaling Blocks

‘RAS Block’: RAS, EGFR, P13K etc

‘TP53 Block’: TP53, Rb, TP14, BRCA1/2



breast

CNS

blood

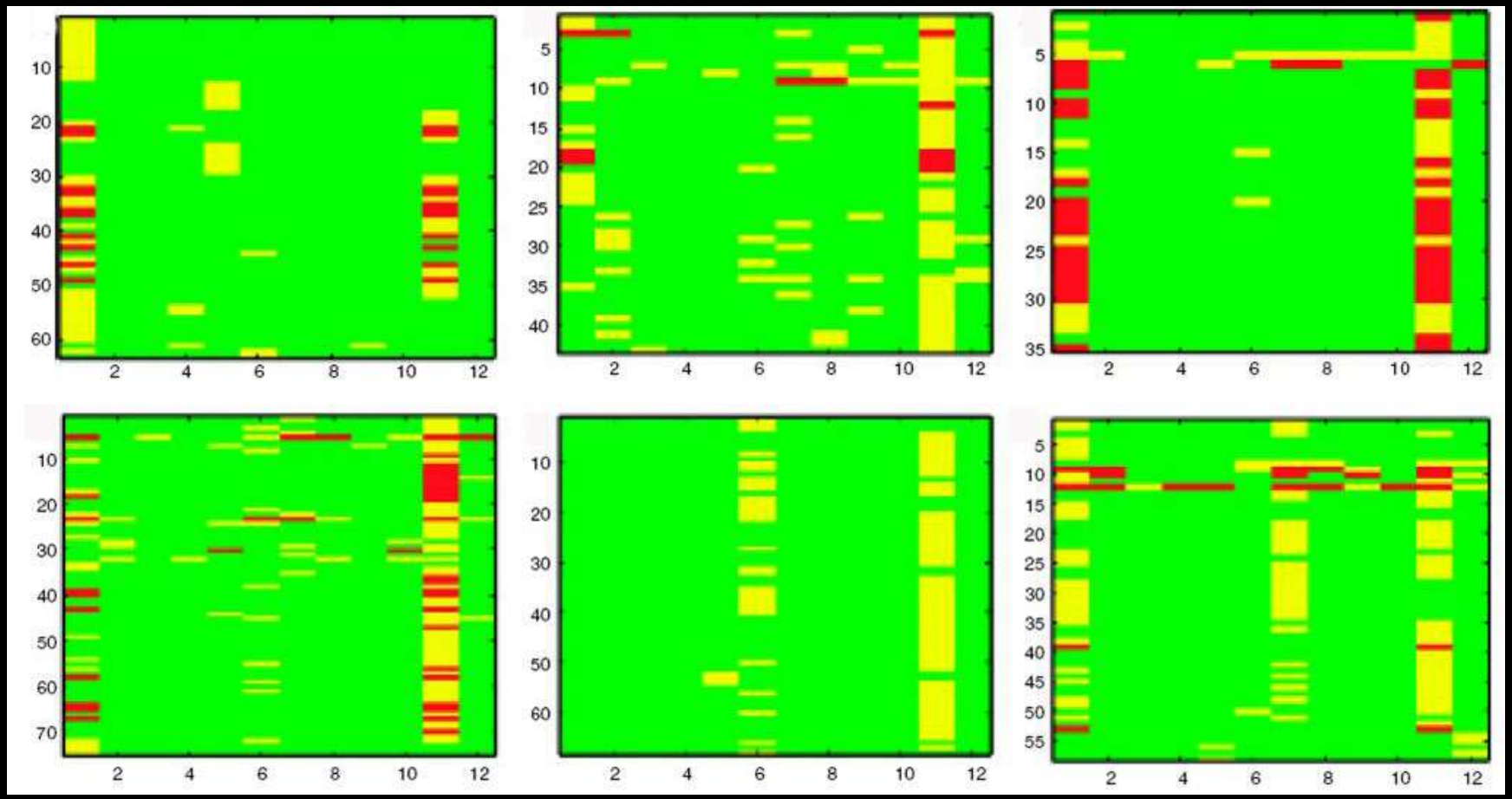
lung

pancreas

skin

Modified from: Q. Cui et al (2007) Mol. Sys. Biol. 3, 152

Heatmaps of Gene Mutation Distributions in 12 Oncogene-Signaling Blocks in Human Cancers



Modified from: Q. Cui et al (2007) Mol. Sys. Biol. 3, 152

ID of Low Frequency Drug-Sensitizing Genotypes Outside of Expected Histologically Defined Cancer Types

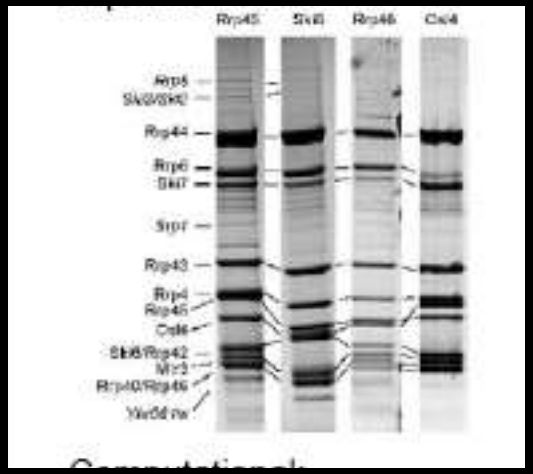
(U. McDermott et. al. 2007 PNAS 104, 19936)

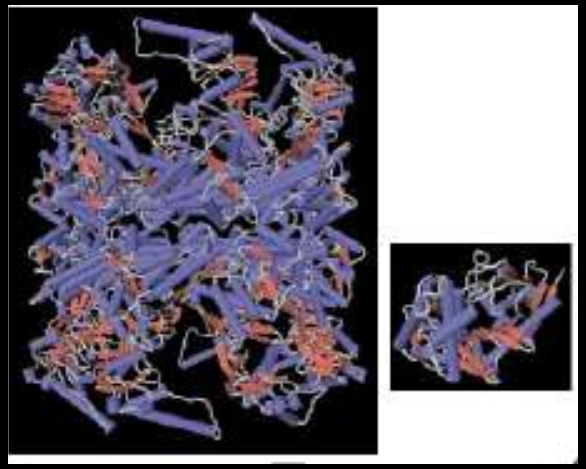
- MET amplification
  - gastric plus esophageal and lung
- HER2 amplification
  - breast plus esophageal, lung, gastric
- BRAF (V600E mutation)
  - melanoma plus thyroid, colorectal
- EGFR
  - NSCLC, esophageal, gastric, ovarian

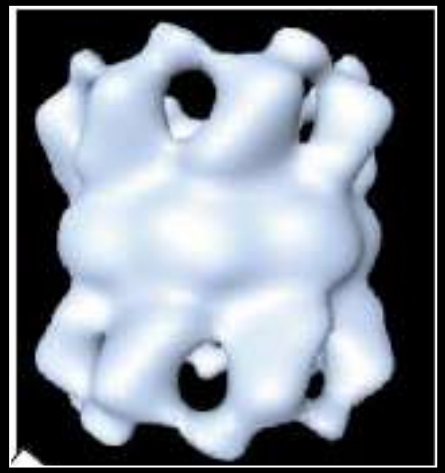
Signaling Pathways in Human Cancers

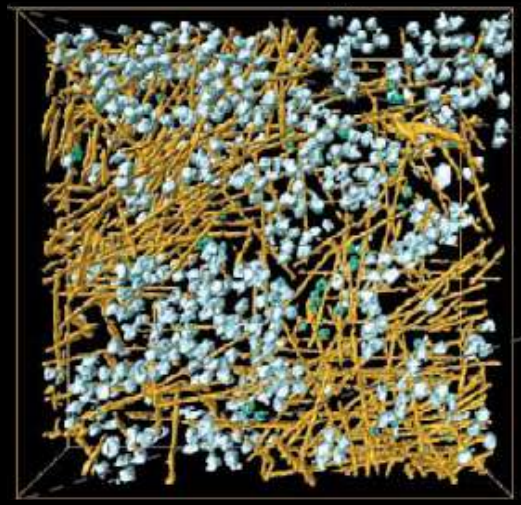
- high frequency of mutations in TP53 block combined with other (activating) blocks (eg. RAS)
- 'TP53 block': loss of function mutations and/or methylation in cell cycle, apoptosis and DNA damage checkpoints genes
  - key event in early tumorigenesis and cancer stem cell maintenance (?)
- activation pathway lesions typically involve gain of function mutations
  - promote cell proliferation and survival

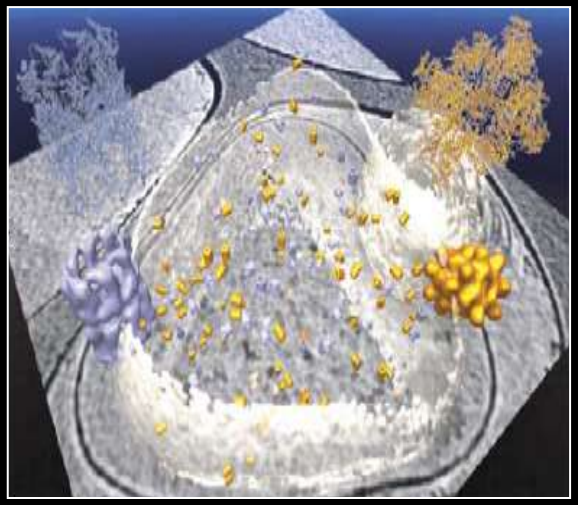
Mapping the Structural and Topological Dynamics of Networks: The Next Grand Challenge

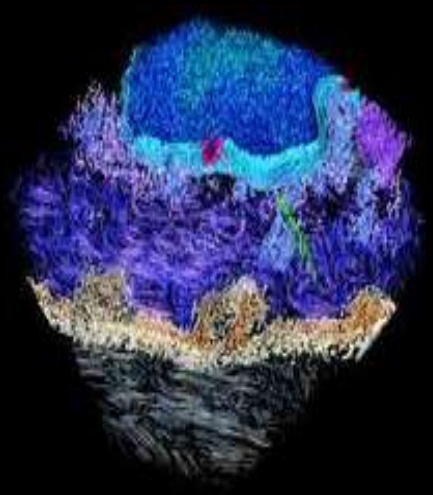






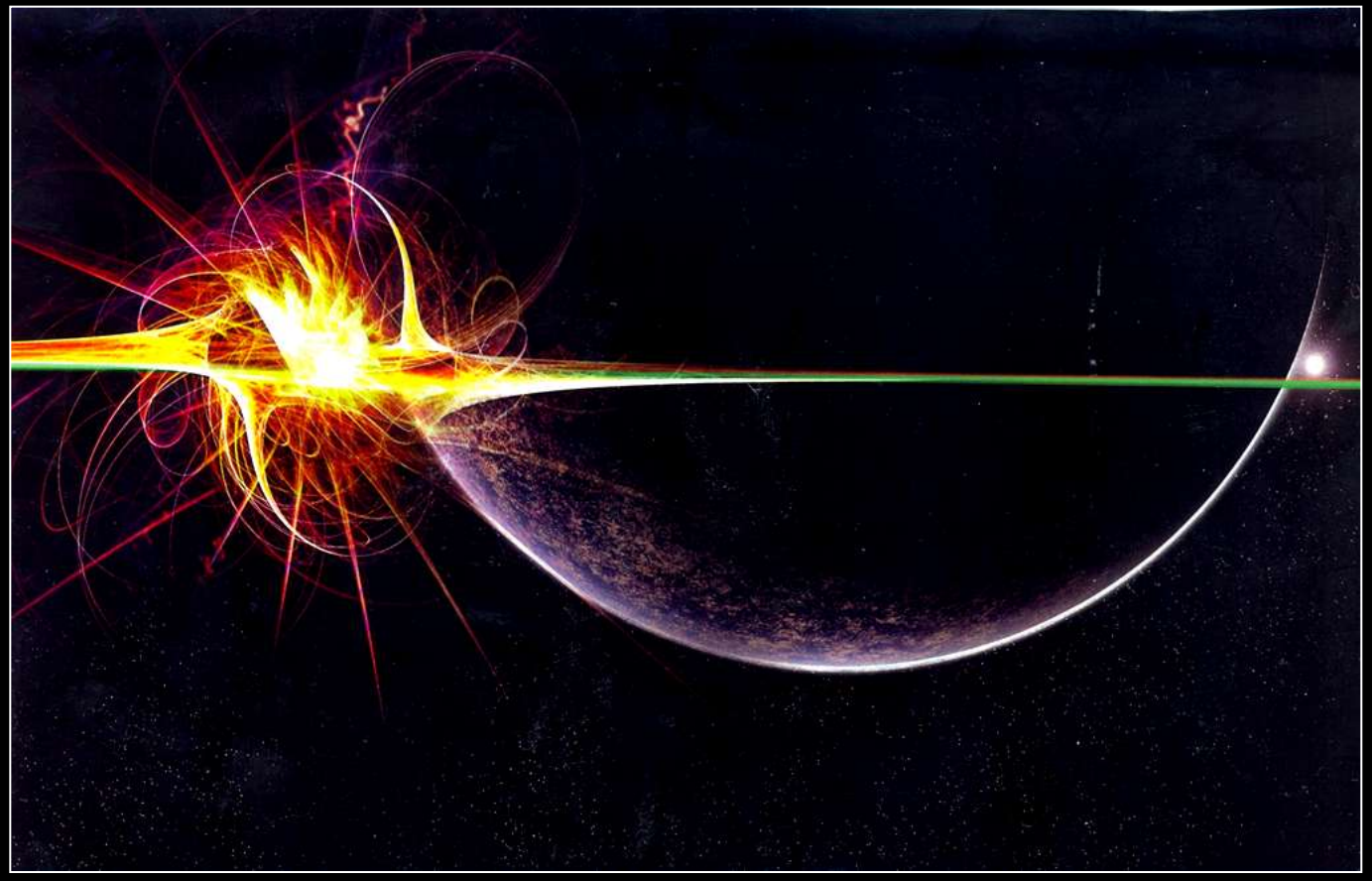






From EMBL: A. Frangakis

The Behavior of Far-From-Equilibrium Systems: Thermodynamics, Genomic (In)Stability and Cancer



The Behavior of Far-From-Equilibrium Systems: Thermodynamics, Genomic (In)Stability and Cancer

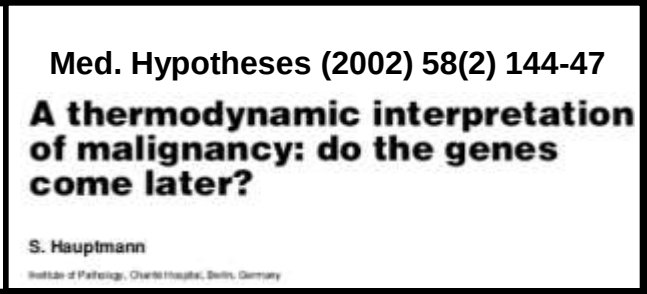
- biological information conforms to the second law of thermodynamics
- informational entropy of the genome will increase over time
  - random loss of molecular fidelity
  - accelerated entropy via environmental stress/insult
- thresholds for fidelity maintenance systems, aging and age-associated diseases
- environmental stress(es) / insult(s) lead to increased cellular entropy via localized energetic overload
  - ID compensatory mechanisms to mitigate entropic decay

A Thermodynamic Interpretation of Malignancy

PLoS Comp. Biol. (2008) 4, 1

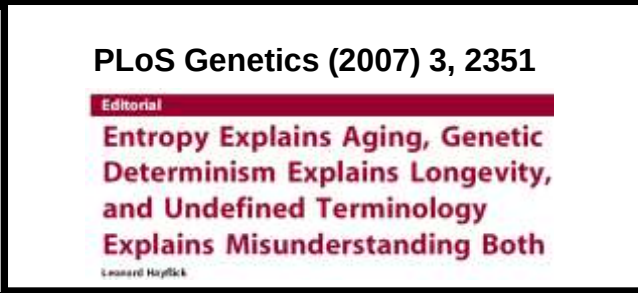
Entropy Measures Quantify Global Splicing Disorders in Cancer

William Ritchie[1], Samuel Granjeaud[1], Denis Puthier[1], Daniel Gautheret[2]*


Med. Hypotheses (2002) 58(2) 144-47

A thermodynamic interpretation of malignancy: do the genes come later?

S. Hauptmann


PLoS Genetics (2007) 3, 2351

Editorial

Entropy Explains Aging, Genetic Determinism Explains Longevity, and Undefined Terminology Explains Misunderstanding Both

- cancer as an adaptive phenomenon to environmental stress(es) / insult(s) causing increased cellular entropy
- response to harmful, localized energetic overload
- genomic instabilities as adaptive response(s) to entrap energy and reduce entropy
- cancer as dissipative structure harmful to the individual organism but the underlying genomic repertoire intrinsic to life's capacity to adapt and to evolve

The Behavior of Far-From-Equilibrium Systems: Thermodynamics, Genomic (In)Stability and Cancer

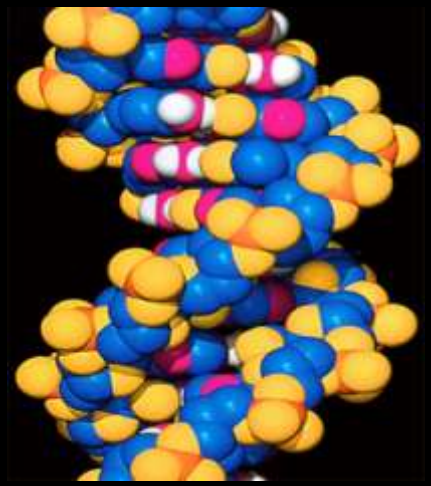
- process of adaptation to insults causing increased entropy in cellular networks
  - mutation and multiplication of DNA
  - aneuploidy
  - alternative protein isoform expression via alternative splicing
- increased synthesis of structural elements with higher intrinsic energy content
  - heat shock proteins
  - altered protein chirality
  - D-amino acids
  - cheater versus cooperator phenotypes (cf. bacterial clone fitness in adverse environments)

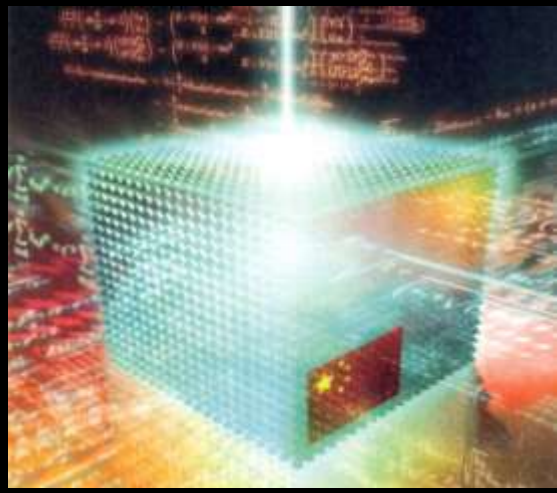
"Managing Mega-Data": (Who Knows Wins)

volume

scale

global networks

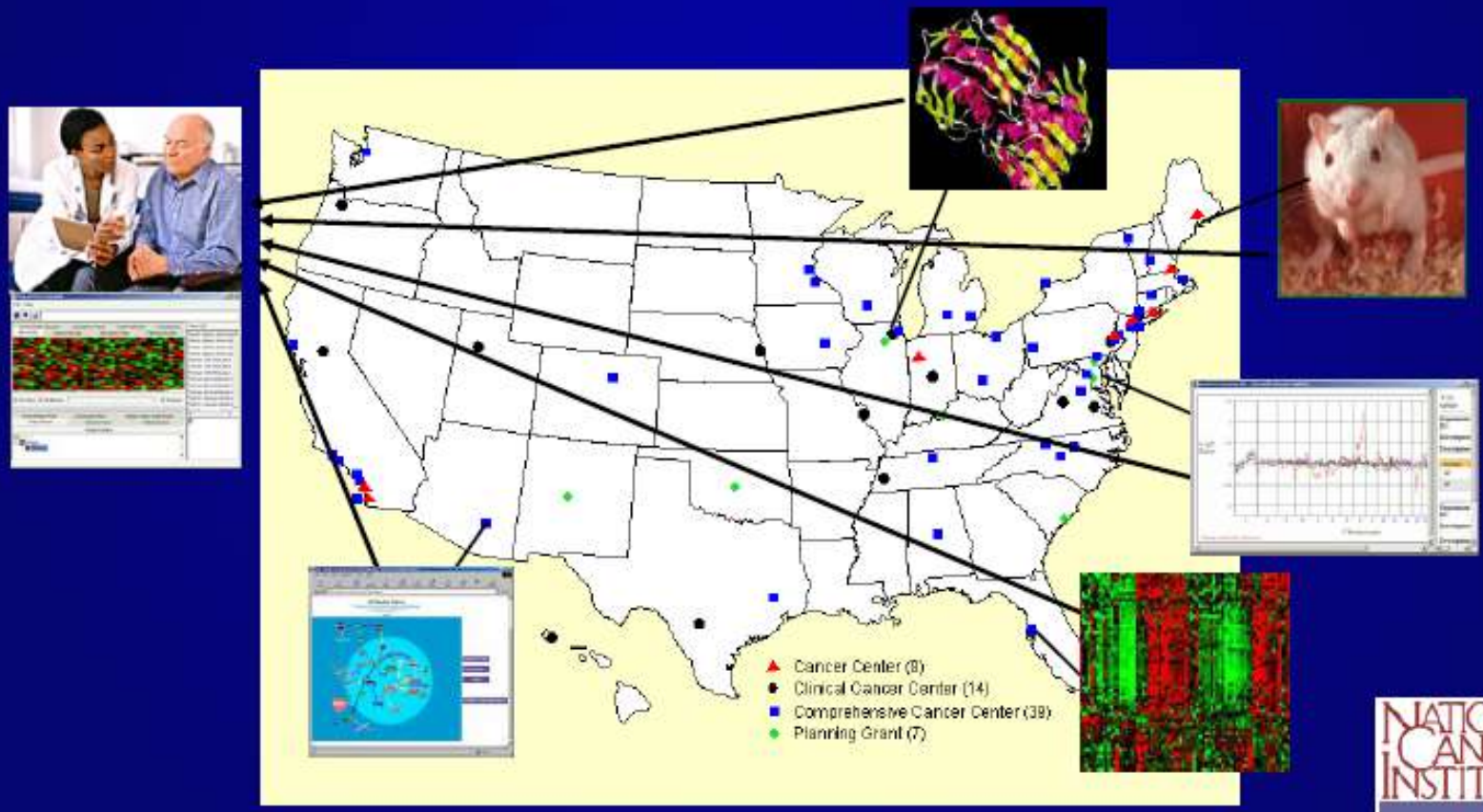




heterogeneity

integration

Taming the Biomedical Data/Information/Knowledge "Beast" – The NCI's Cancer Bioinformatics Grid



Systems Biology, Molecular Profiling of Disease and Personalized Medicine

A Realistic Strategy for the Delivery of Rational Healthcare?
Or
An Erstwhile Intellectual Pursuit Doomed to be Dashed on the Rocks of Scientific Silos, Clinical Conservatism or Commercial and Regulatory Myopia?

Reimbursement for Diagnostic Tests

- inadequate US Medicare coding and payment mechanisms
  - fixed, out moded, out-dated, lacking in transparency, inconsistently applied
- no effort to link reimbursement to value
- inappropriate assignment of existing CPT codes to new tests
- engagement of third party payers who derive economic/clinical value from new Dx
  - Genomic Health Oncotype Dx

Pricing barriers to Adoption of Molecular Diagnostics for Personalized Medicine : Oncology

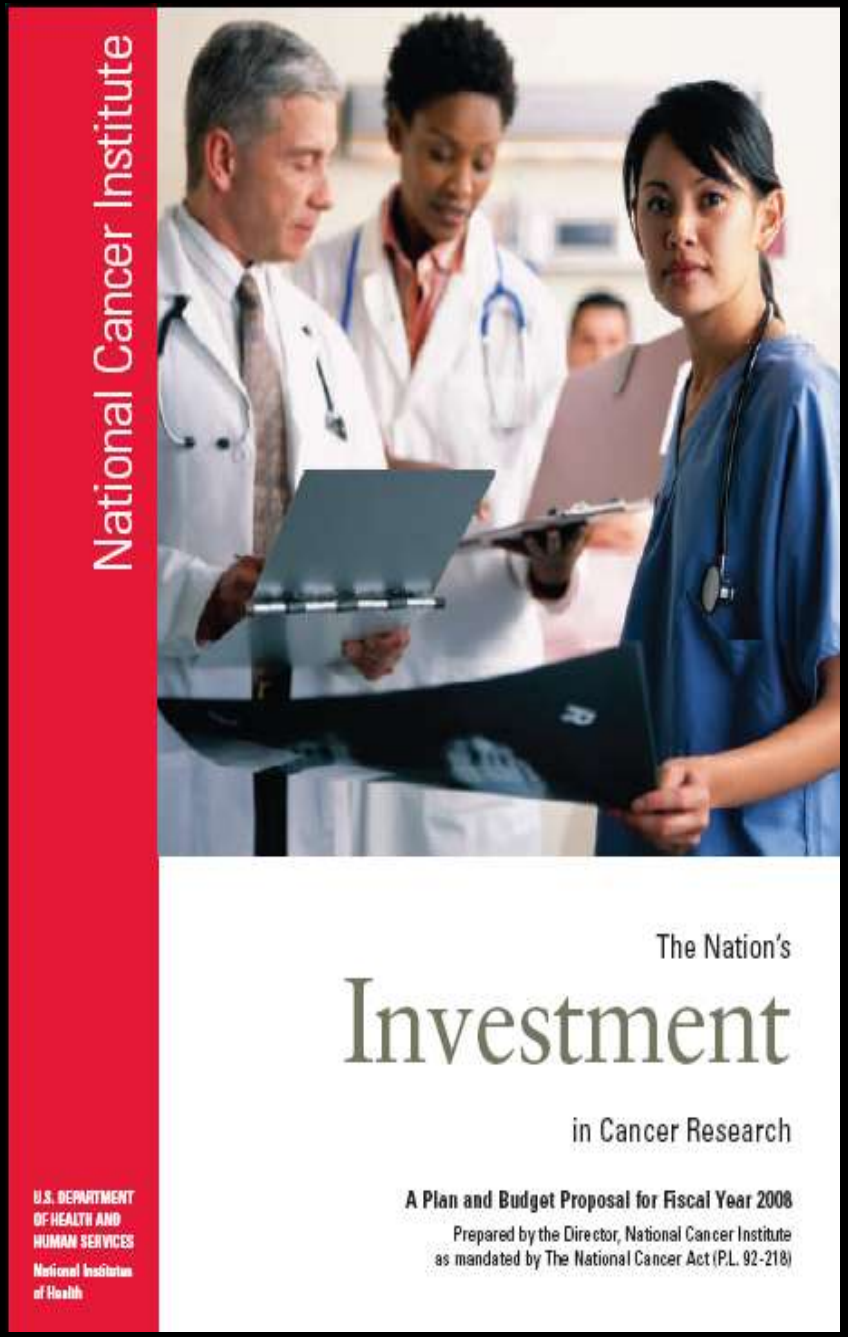
- 1.4 million cancer diagnoses per year (USA)
- lung, breast, prostate and colon account for 55% of all cases
  - 155-234,000 cases for each malignancy
  - all other cancers = < 65,000 cases per year
  - pediatric malignancies (0-15 years) = 9000 cases
  - adoption/pricing for recovery of $100 million R&D costs (7 year market life)
    - $100 test = 142,000 pts; $500 test = 28,000 pts

Comprehending Networks in Complex Adaptive Biological Systems:

- the Intellectual foundation of molecular medicine and rational therapeutics
- mapping network perturbations in major diseases
  - cellular specificity and network design
  - disease-associated ‘biosignatures’ of network disruption
  - ‘signatures’ for disease predisposition
- novel multiplex molecular diagnostics
  - accurate diagnosis and disease subtyping
  - early detection and optimum Rx selection
- network pharmacology
  - design of targeted Rx
  - customized selection of optimum Rx
- remote monitoring of health status and Rx compliance
  - POC, ‘Doc-in-a-Box’

Mastery of Large Scale Biology and Daunting Technical and Logistical Complexities

- obligate embrace of technical convergence
  - life sciences, engineering, computing
  - novel cross-sector relationships
- scale
  - escalating cost and new organizational structures/alliances
- high resolution, high throughput analytical tools
  - real-time ‘global profiling’ of biological network dynamics
- the imperative for comprehensive standardization of analytical data
- integration of massive, heterogeneous datasets
  - new mathematical and computational algorithms
  - information architectures and datamining capabilities
- increasing use of consortia and open-source resources



Strategic Imperatives

- convergence
- integrated approaches
- connectivity
- leveraging resources and knowledge
- community

Major Technology Needs

- low cost whole genome sequencing ($1000 human genome)
- rapid gene ID in complex genomes
- structural genomics and protein structure-function prediction
- mapping network dynamics
  - real-time global mapping of gene expression, regulatory controls and proteome dynamics

Creating New Networks of Connected Expertise to Accelerate Innovation in Healthcare R&D

ASU PROPRIETARY

From Silos to Systems-Based Strategies

- extravagant waste of uncoordinated, fragmented research
- fragmentation reinforced by anachronistic government funding policies
- insufficient interdisciplinary leverage of convergent technologies (academia and industry)
- inadequate standards for molecular profiling data
- systemic deficits in electronic connectivity in healthcare as major obstacle in integration of molecular profiling with disease patterns and treatment outcomes
- inadequate market incentives for integration of Dx, Rx and Ix products/services and healthcare delivery

ASU PROPRIETARY

The Imperative for New Approaches: The Launch of Important New Strategic Programs

Development of New Protein Capture Technologies

NCBI
Entrez, The Life Sciences Search Engine
ToPaz
caBIG cancer Biomedical Informatics Grid
International HapMap Project
PLoS one BETA
Publishing science, accelerating research
ALLEN INSTITUTE for BRAIN SCIENCE | Allen Brain Atlas
W3C WORLD WIDE WEB consortium
PubMed www.pubmed.gov
FDA/Severe Adverse Events (SAE) Consortium
ProteomeBinders
science commons
The Neurocommons
THE CANCER GENOME ATLAS
Welcome to HuGENet
BROAD INSTITUTE
Diabetes Genetics Initiative
CRITICAL PATH INSTITUTE
Improving the Path for Innovative Therapies
PGRN
creative commons
Foundation
THE biomarkers CONSORTIUM
NATIONAL INSTITUTES OF HEALTH
Genes, Environment and Health Initiative (GEI)
Determining Genetic and Environmental Roots of Common Diseases
Clinical Semantics Group

The Increasing Importance of Open Source Resources in Life Sciences R&D

- consortia
  - scale
  - share risk and cost
  - standards
- open-source standards are not incompatible with IP protection and commercial incentives
- new networks of turnkey services

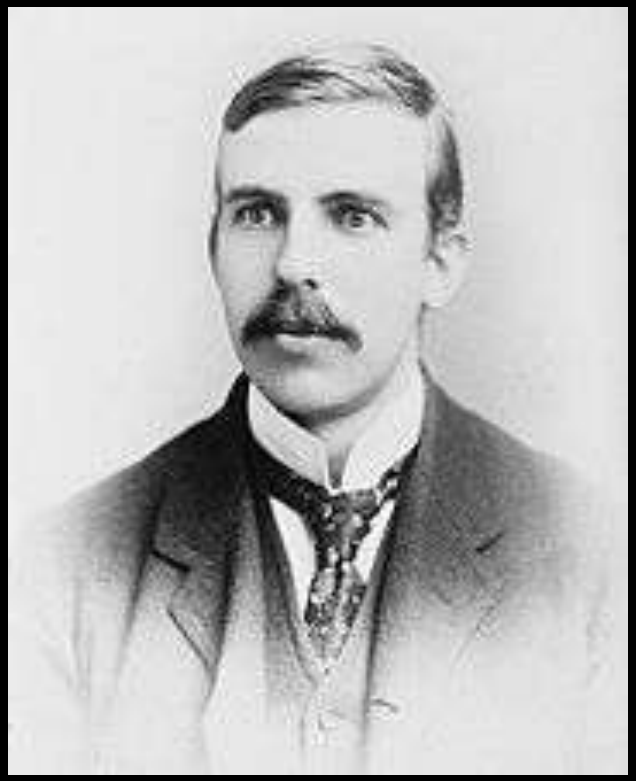
Health Technology Networks: Leveraging Expanding Information Content, Tools and Services

- ever faster generation of new information
- current R&D ecosystem is too fragmented to fully leverage novel content and shared learning
- rise of new business models of ‘expertise networks’ that eclipse current monolithic single company innovation models

Elucidation of the Design Principles of Biology Complexity

Era One: "It from Bits"

- physiology and cell biology
- molecular biology
- systems biology
  - mapping information coding and flow in complex adaptive systems
  - defining the 'rule sets' of biological assembly and order
- learning from precedence/predictive modeling of scale-free networks in non-medical domains
  - military
  - climate
  - telecommunications
  - global finance



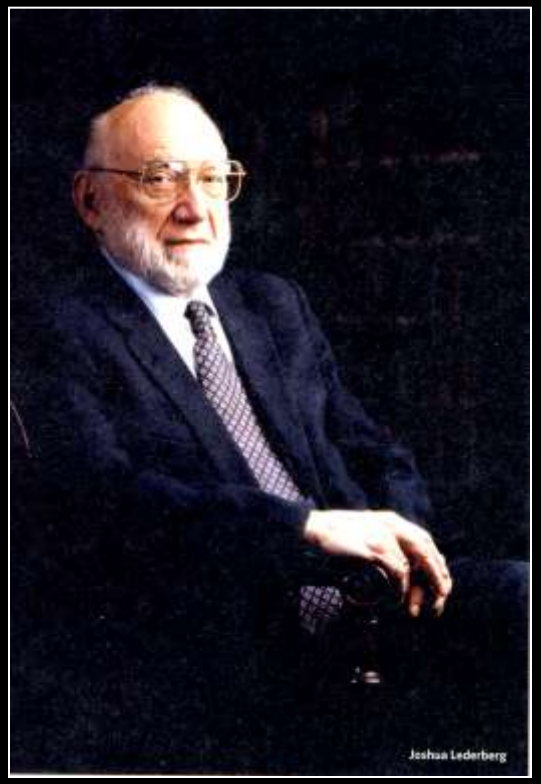
All science is either physics or stamp collecting. (Ernest. Rutherford)

Meeting Report

Integrating and Leveraging the Physical Sciences to Open a New Frontier in Oncology

February 26-28, 2008 | The Ritz-Carlton | Pentagon City | Arlington, VA

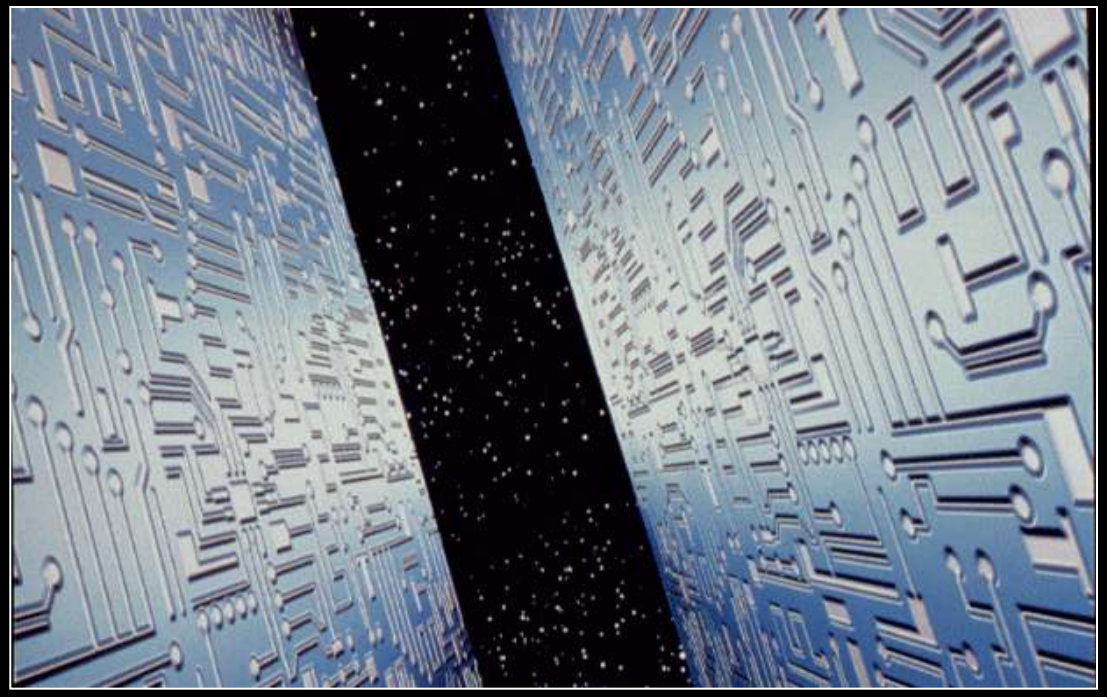
Visionaries



J. Lederberg

A. C. Clarke

Aspiration and Engagement with Grand Challenges



"The only way of discovering
the limits of the possible
is to venture a little way past them
into the impossible"

Arthur C. Clarke
Profiles of the Future (1962)