Systems Biology and Systems Medicine

Lee Hood

Institute for Systems Biology, Seattle

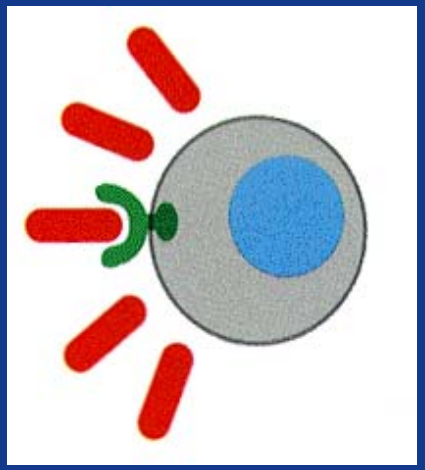
Biology is an Informational Science

There are two types of Biological Information

- The digital information of the genome

- The environmental information that impinges upon and modifies the digital information.

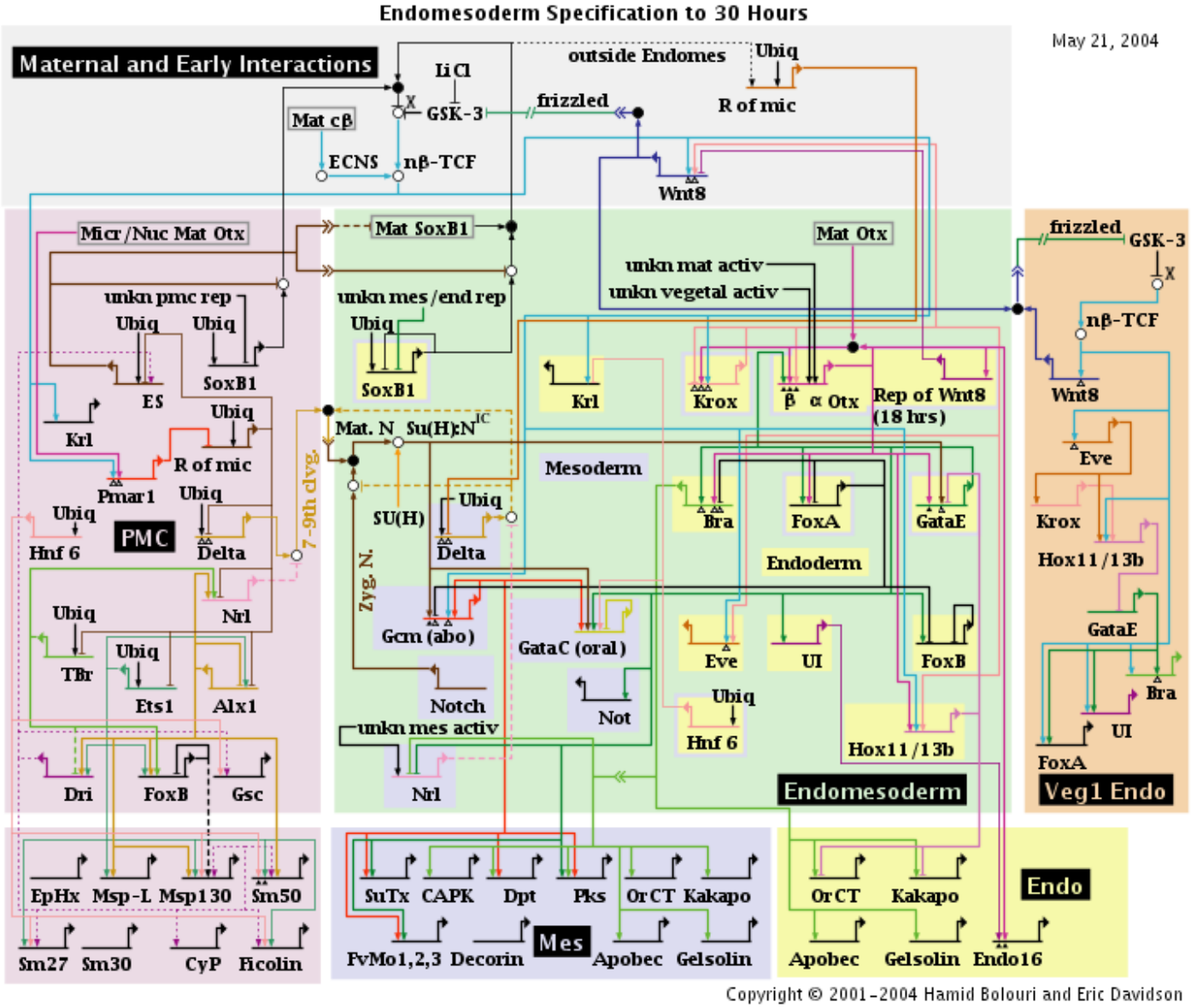
```
CCAGAAAGGC CGAGGCTCTG CAGCGGGAGG
GCAGGGCACA GGGACAGCCC CCCTCCACAG
CCAGGAGGTT GCTTCTTCCA GGAGGCTTTT
GCTCCCAGCT GCTGTGAGTG CTGCACATTC
CACTTCTGGT GCCCACTGTG GCCACAGCAA
GCCTCCTGGG GAGCTGCTGA CCCTAGGCAG
CACCCCAGTG TTTGCCAGTG TTTGCCCGTG
TTTGCTCGCC AGTGTTCGCC ACTTGTCCCT
GAAGTTGCAG GTCCCTCCAG GACAGTTGGC
```



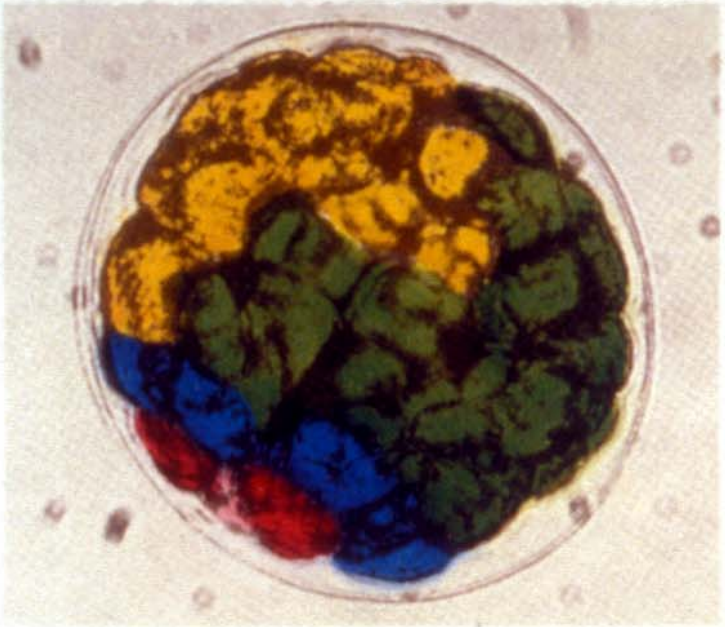
Two Types of Systems Biology

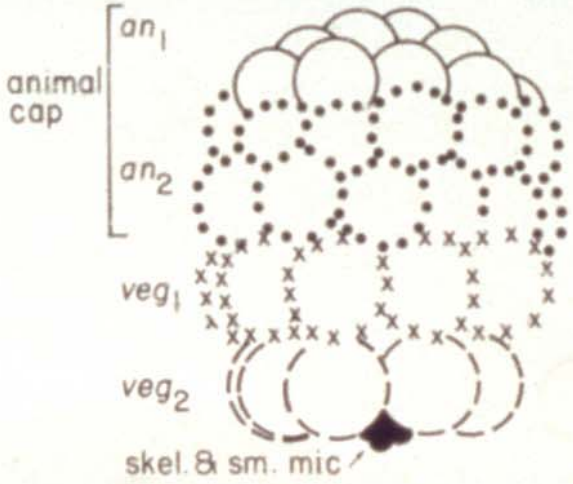
- Decipher function of molecular machines--execute biological functions
- Decipher function of biological networks--capture, transmit, integrate and disperse biological information
- Two types of biological networks: protein networks and gene regulatory networks

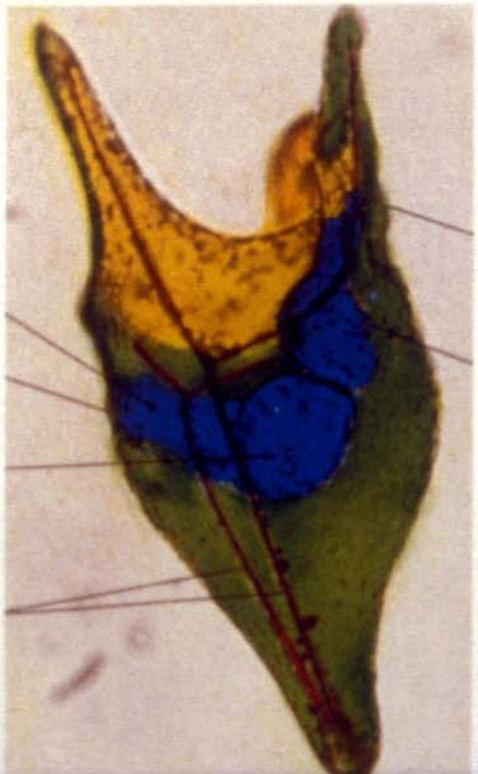
Most Sophisticated Integrated Biological Network Defined to Date



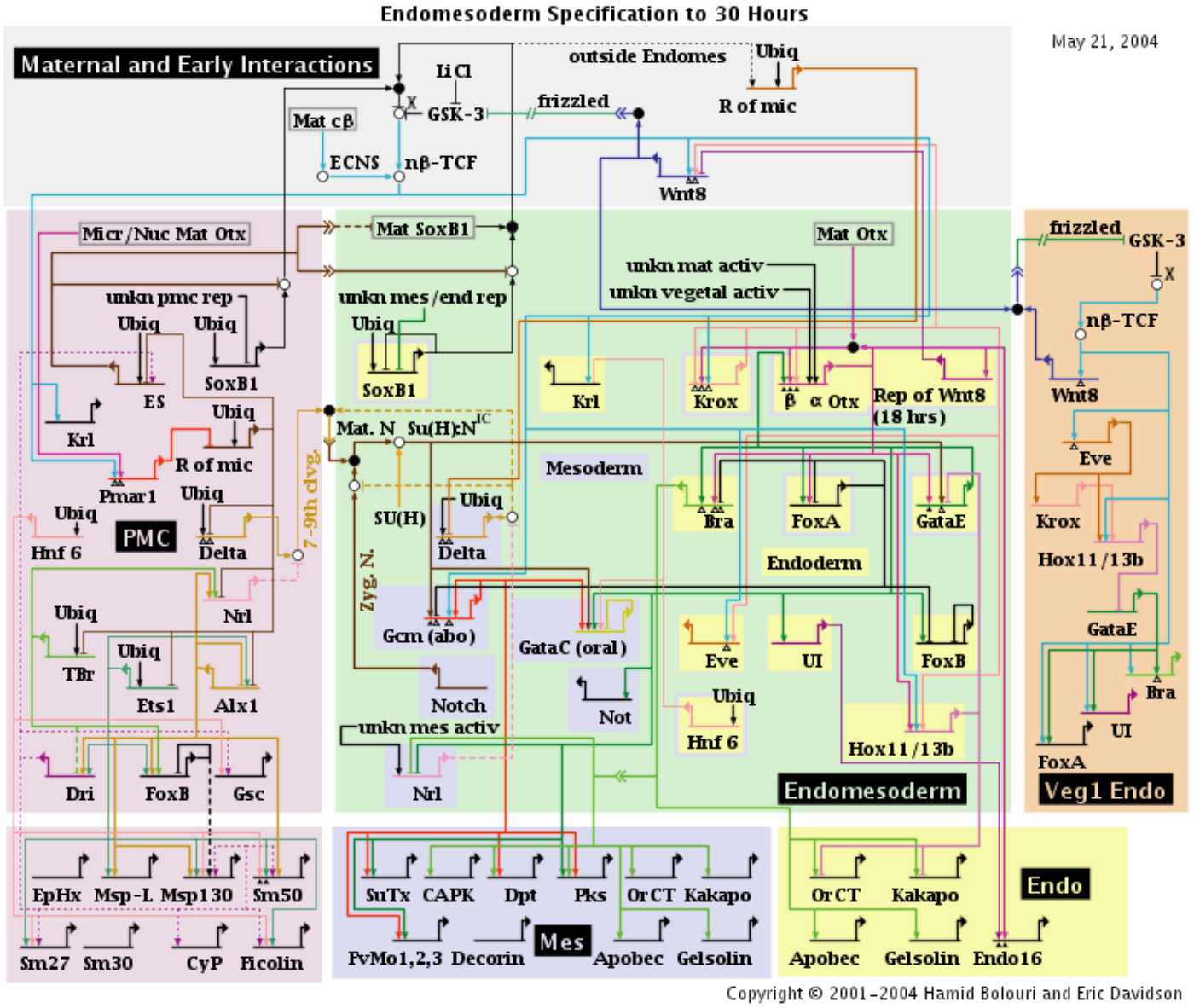
Sea Urchin Larval Development







Most Sophisticated Integrated Biological Network Defined to Date



Networks Process Biological Information: Transmit, Integrate, Execute

Institute for Systems Biology

Hierarchical or Multiscalar Levels of Biological Information

Top down, bottom up and both ways from middle

Level of System Analysis

Connect to Digital Core

Integration of Different Levels

DNA

mRNA

Protein

Protein interactions and biomodules

Protein and gene networks

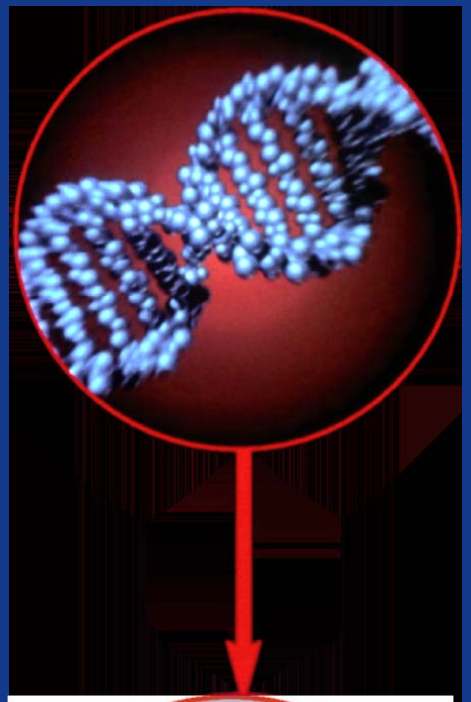
Cells

Organs

Individuals

Populations

Ecologies



Institute for Systems Biology

Systems Biology at ISB

INSTITUTE FOR Systems Biology

Technology Has Transformed Contemporary Systems Biology

Six essential features of contemporary systems biology

***Quantitative measurements* for all types of biological information.**

***Global measurements*--measure dynamic changes in all genes, mRNAs, proteins, etc, across state changes.**

Computational and mathematically *integrate different data types*--DNA, RNA, Protein, Interactions, etc.--to capture distinct types of environmental information.

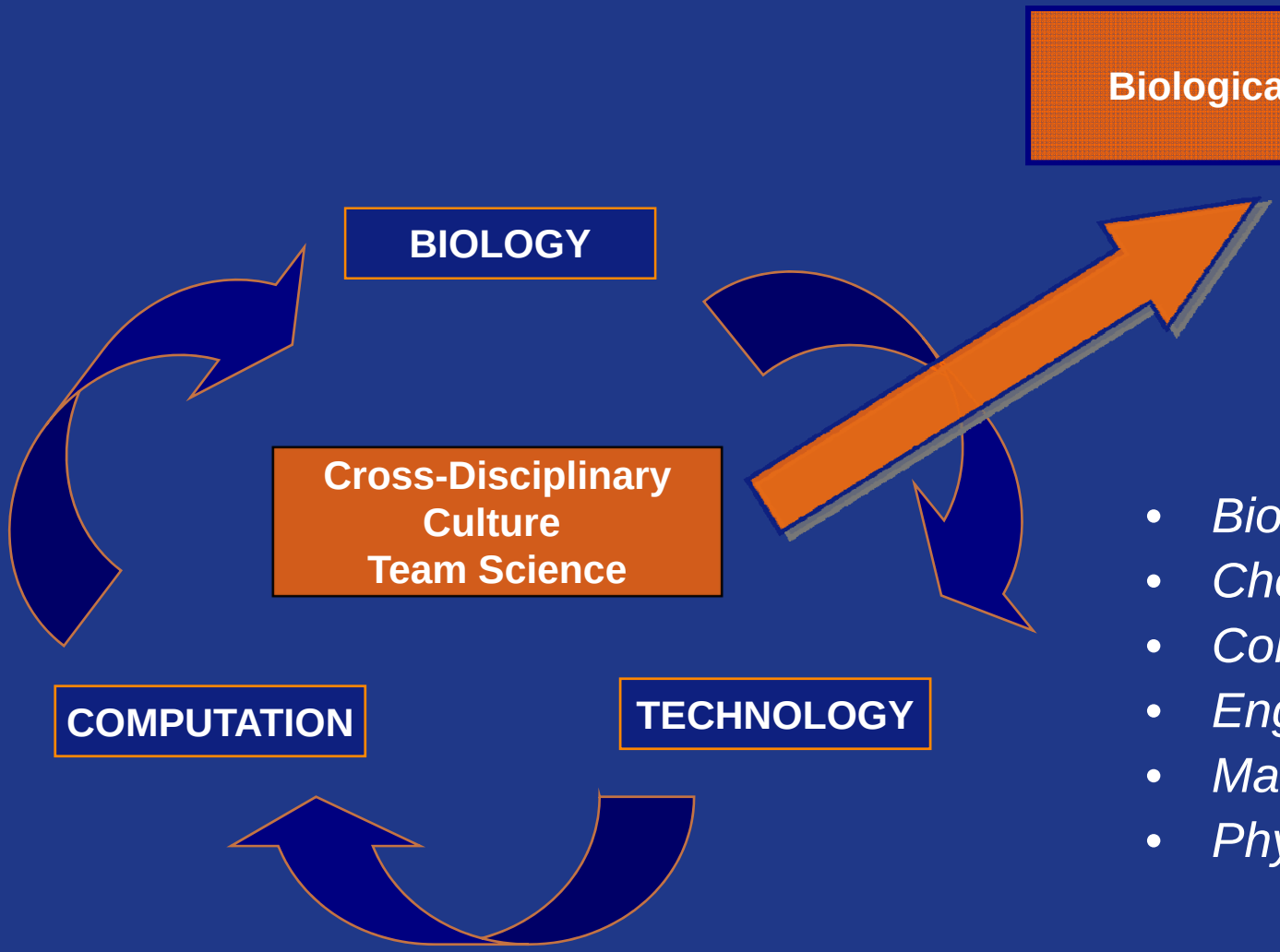
***Dynamic measurements*--across developmental, physiological, disease, or environmental exposure transitions.**

Utilization of carefully formulated *systems perturbations*.

Integration of *discovery-driven* and *hypothesis-driven* (global or focused) measurements . The systems biology cycle: perturbation--measurement--model--hypothesis--perturbation--etc.

INSTITUTE FOR Systems Biology

Agenda: Use biology to drive technology and computation. Need to create a cross-disciplinary culture.



- *Biology*
- *Chemistry*
- *Computer Science*
- *Engineering*
- *Mathematics*
- *Physics*

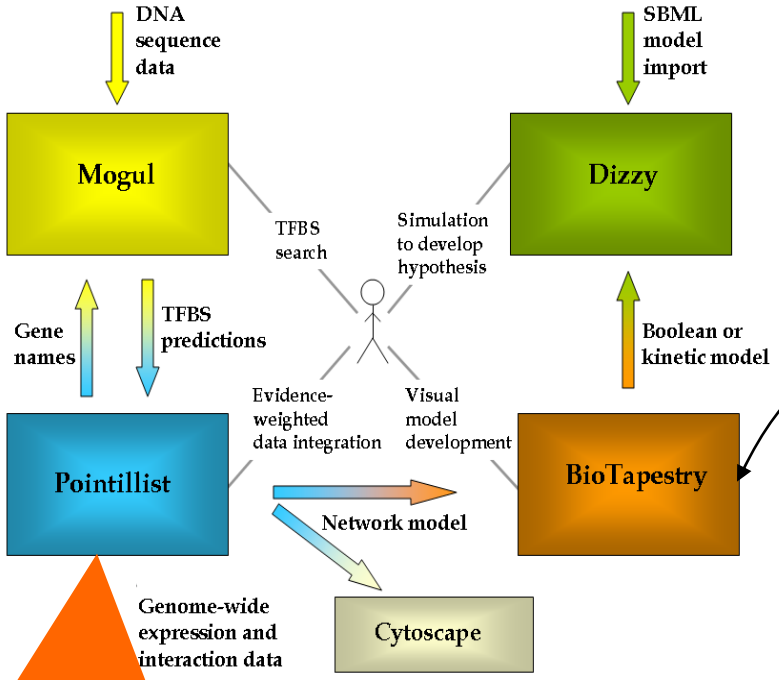
ISB Technology Development

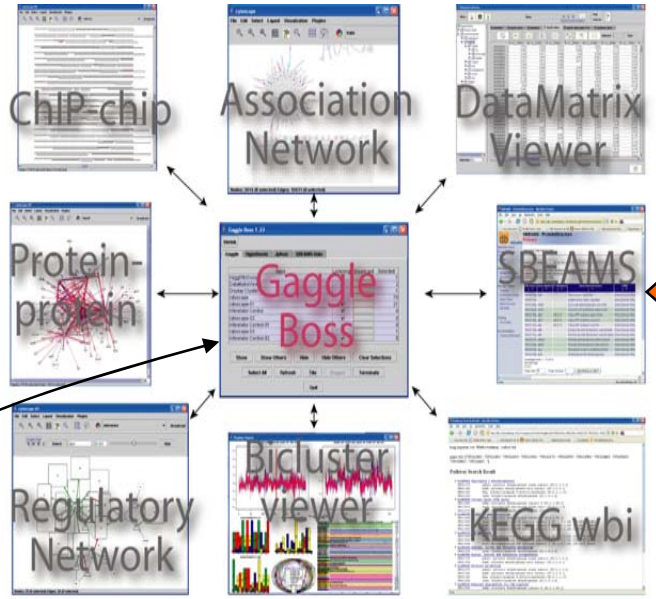
- Genomics--Hood
  - Single strand DNA sequencing--Helicos
  - Protein folding and annotation--Hood/Baliga
  - Gene finding--FEAST
  - Personalized genome analysis--George Church--Harvard
- Proteomics-
  - Analyzing molecular machines--Ranish
  - Glycocapture of blood proteins--Aebersold/Hood
  - Serum databases--reanalysis of primary data--Atlas--Aebersold
  - Computational software programs for validation--Aebersold
  - Toward global proteomics--Hood/Aebersold
  - Measuring 2000 blood proteins--nanotechnology--Hood/Heath
  - Surface plasmon resonance--Hood, Lausted
  - Nanotechnology and microfluidic measurements of protein properties
- Single cell analyses--Ozinsky/Galitski/Aitchison/Hood/Heath
  - Microfluidics
  - Nanotechnology
- Imaging and visualization---strategic partnerships

Techniques for searching new realms of data space

Informatics at ISB

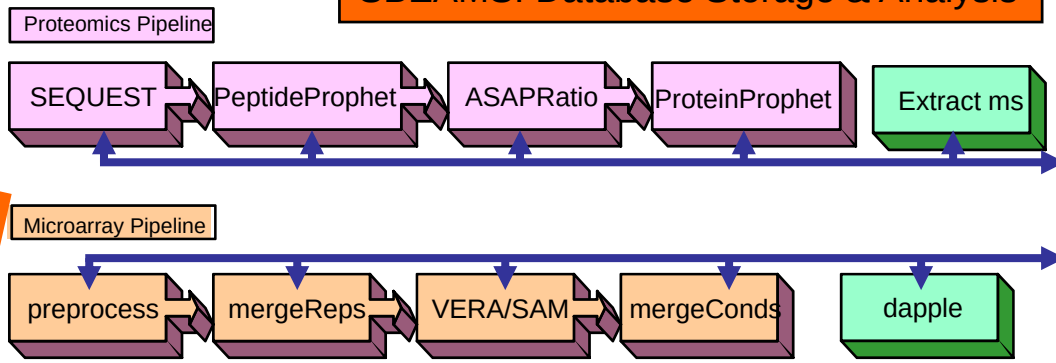
Model Refinement & Analysis Tools

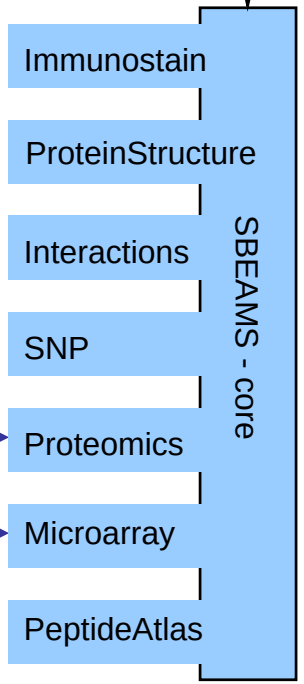




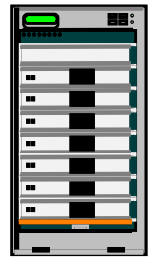
Gaggle: Network & Data Visualization

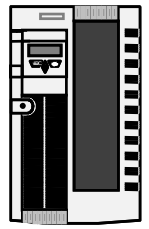
SBEAMS: Database Storage & Analysis





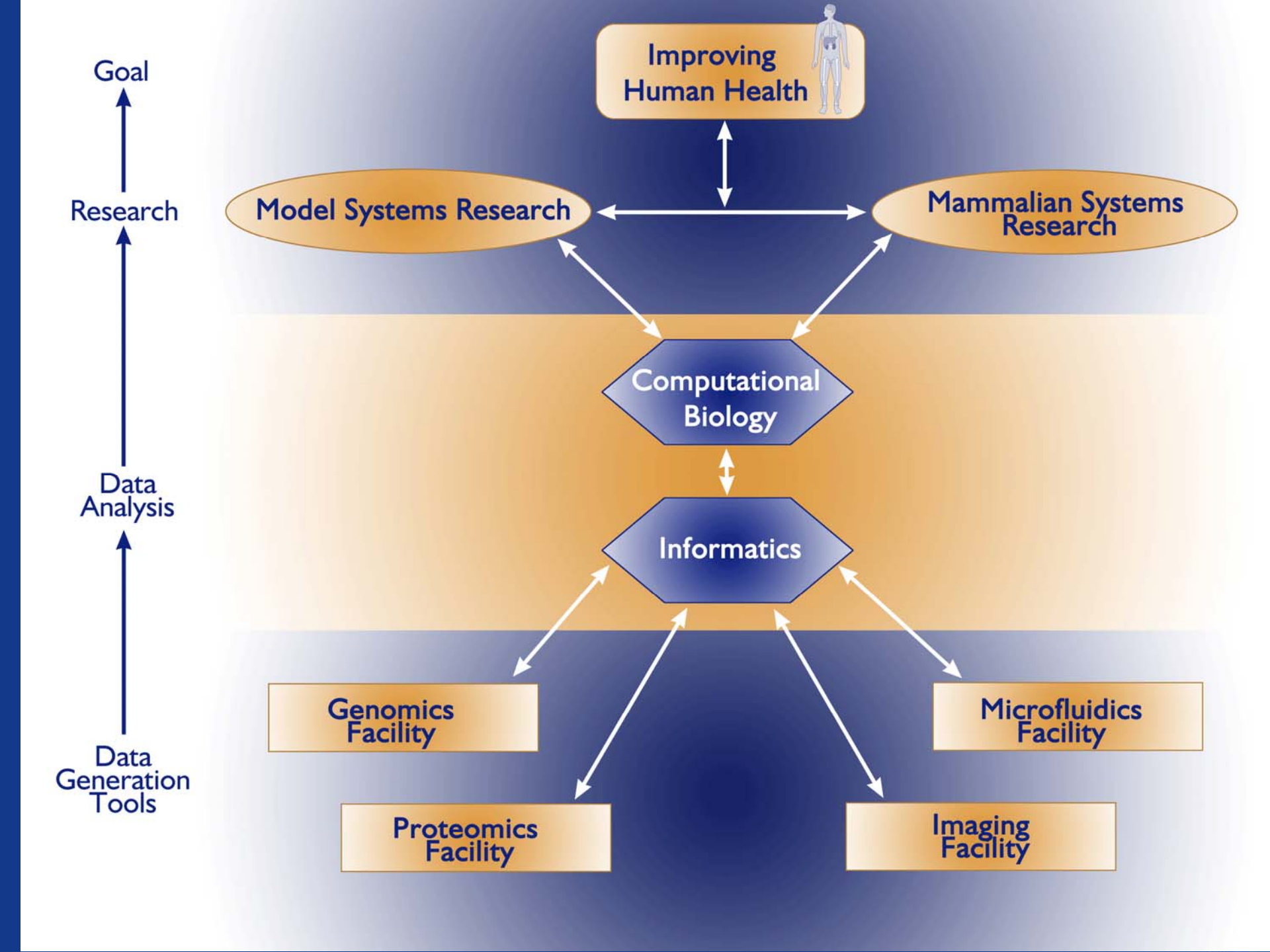
File Servers





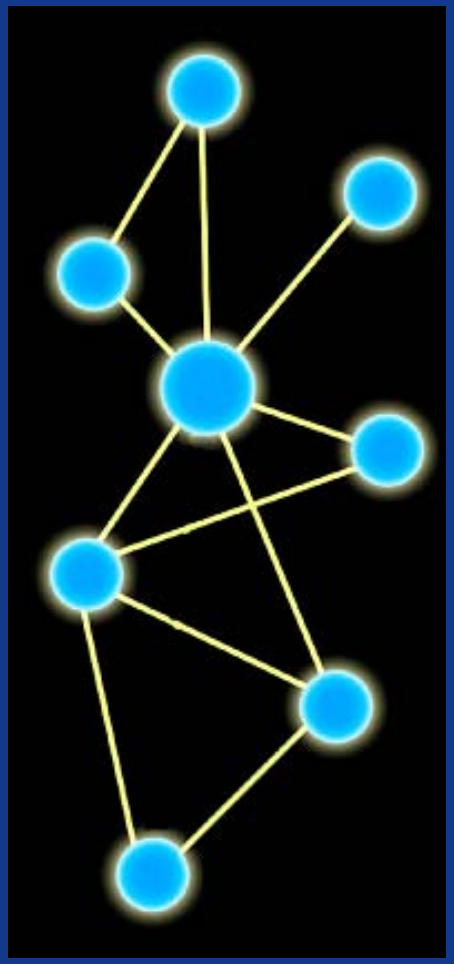
SQL Servers

INSTITUTE FOR Systems Biology

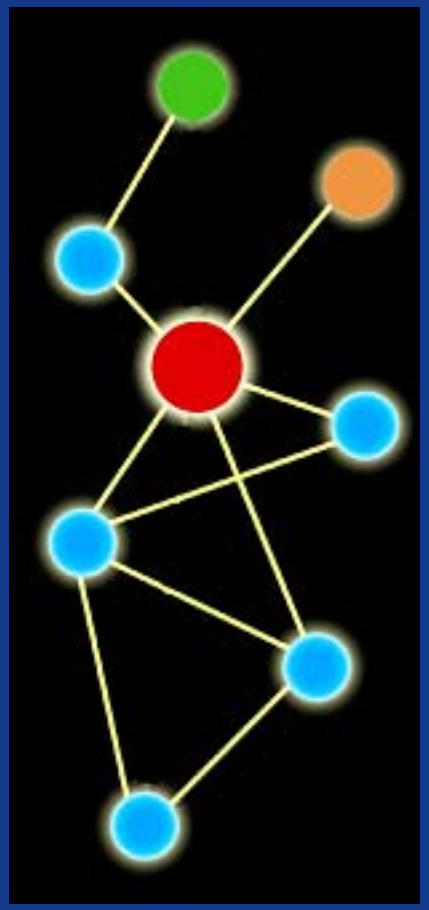

Goal
Research
Data Analysis
Data Generation Tools
Improving Human Health
Model Systems Research
Mammalian Systems Research
Computational Biology
Informatics
Genomics Facility
Microfluidics Facility
Proteomics Facility
Imaging Facility

A systems approach to disease

Disease Arises from Disease Perturbed Networks



Non-Diseased



Diseased

dynamics of pathophysiology

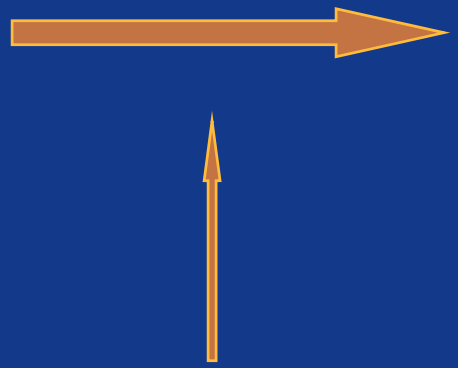
diagnosis

therapy

prevention

Prion Protein Exists in Two Forms

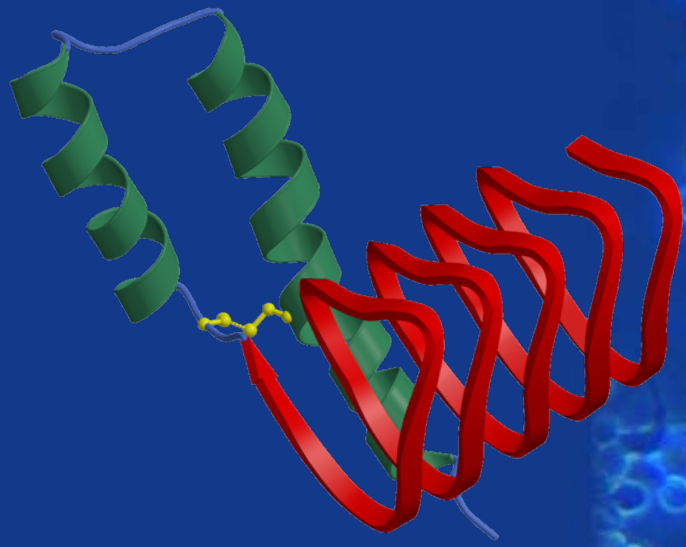
Cellular
PrP^{C}



PrP Genetic Mutations

PrPSc Infections

Spontaneous conversion



Infectious **PrP^{Sc}**

Institute for Systems Biology

Prion Diseases

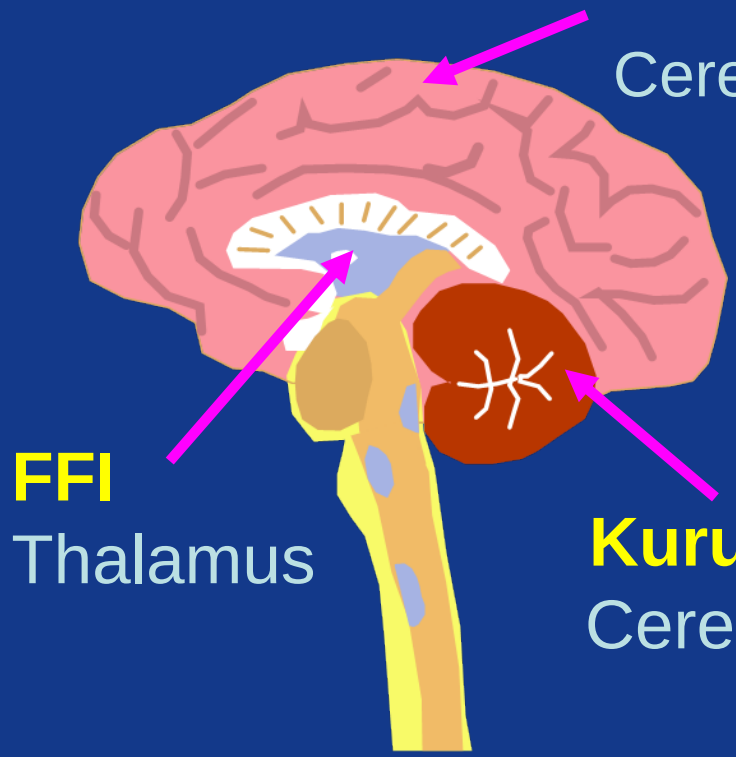
Human:

Kuru of the Fore people in New Guinea

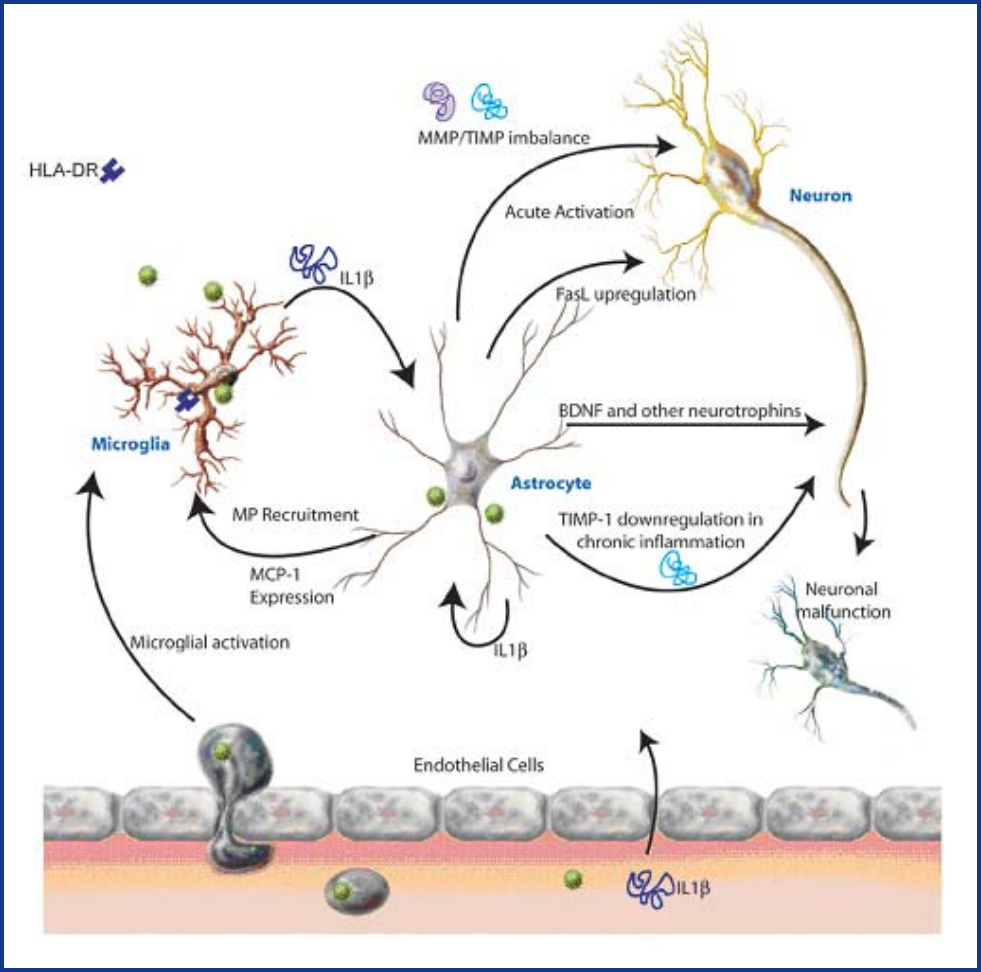
CJD (Creutzhfeldt-Jakob Disease) : Dementia

GSS (Gerstmann-Sträussler- Scheinker) :Ataxia

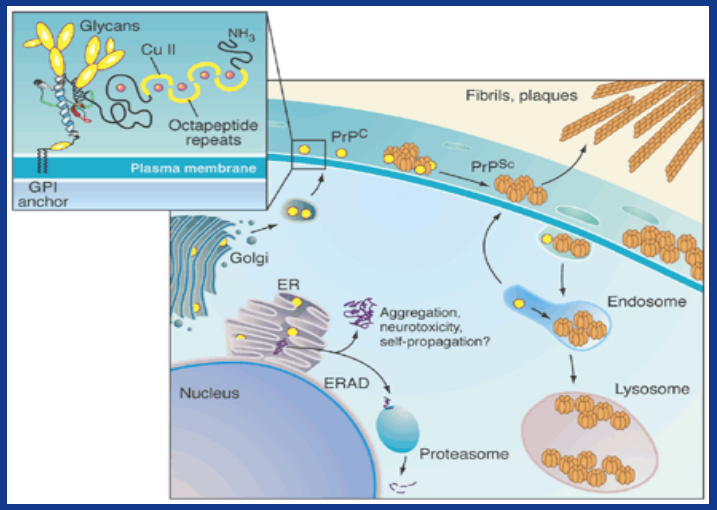
FFI (Fatal Familial Insomnia) : Insomnia



Glial Cells and Neurons

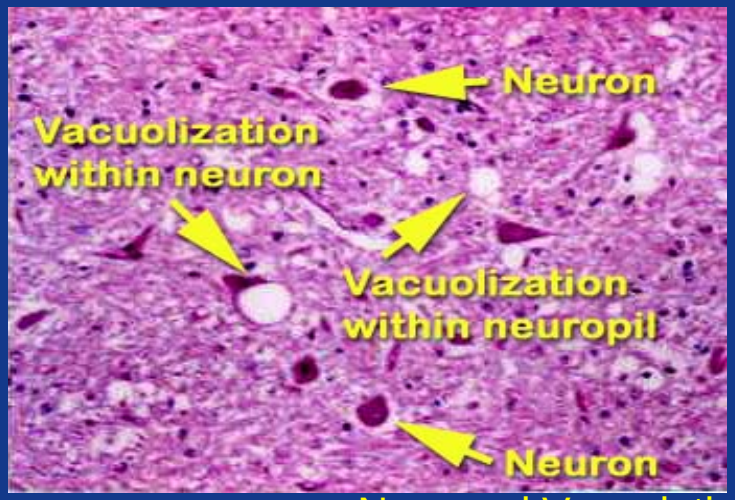


Neuropathological Features



Priola et al., 2003 Science

- PrP accumulation -> Vacuolation
- Microglial/Astrocyte activation
- Synaptic degeneration:
  - Presynaptic bouton degeneration
  - Dendritic atrophy
- Nerve cell death



Neuronal Vacuolation

Multiple groups:five inbred strains, two transgenic strains and one knockout strain

Group	Mouse	*Prnp* Genotype	Prion Strain	Incubation Time (d)
1	C57BL/6J	*a/a*	RML	~150
2	B6.I-1	*b/b*	301V	~120
3	FVB/NCr	*a/a*	RML	~150
4	B6.I-1	*b/b*	RML	~350
5	C57BL/6J	*a/a*	301V	~260
6	(FVB x FVB.129-*Prnp*^{tm1Zrch})	*a/0*	RML	~400
7	Tg(MoPrP-A)B4053	*30 x a*	RML	~60
8	FVB.129-*Prnp*^{tm1Zrch}	*0/0*	RML	No illness

Differentially Expressed Genes--DEGs

INSTITUTE FOR Systems Biology

Systems Analysis of Prion Disease

- Integrate multiple types of data
  - Differential gene expression over time--924 DEGs
    - Brain (and spleen)
    - Two prion strains, seven mouse strains (plus *Prnp*^{tm1Zrch} control)
  - Protein interaction and gene regulatory networks for hypothesis building
  - PrP^{Sc} accumulation and regional deposition over time
  - Integrate with regional differences in gene expression (Allen Brain Institute)
  - Disease symptoms

Differentially Expressed Genes shared by five prion-mouse combinations: Do they encode major disease responses

924 genes shared by five mouse-prion combinations are highly likely to be directly/indirectly involved in prion replication and neuropathogenesis.

Use protein interaction and gene regulatory databases to build hypothetical protein interaction networks for major disease features:

- Microglial activation
- Astrocytic hypertrophy
- Presynaptic bouton degeneration
- Dendritic atrophy
- Prion replication and accumulation
- Nerve cell death

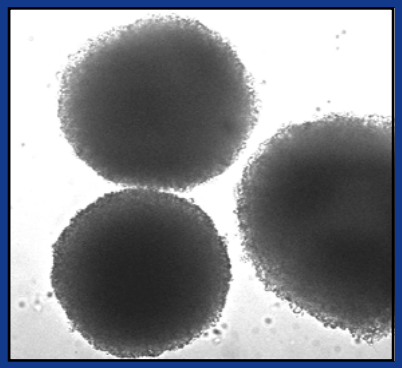
DEGs Encoding Known and Novel Prion Disease Phenotypes

- 7400 Differentially Expressed Genes (DEGs) in 5 inbred strains upon prion perturbation.
- Biological filters reduce to 924 core DEGs for prion disease
- 253/924 DEGs encode known disease phenotypes
- 671/924 DEGs encode novel disease phenotypes

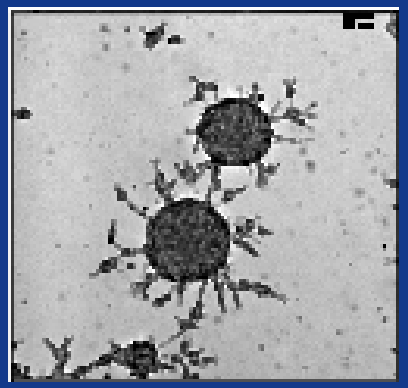
Correct Cellular Assignments of mRNA Expression

- Allen Brain Institute--in situi hybridization of mouse brain with 20,000 genes--localization of transcripts to individual cells
- A perfect match between our assignments of differentially expressed genes (DEGs) in prion disease to brain cell types and those identified by the Allen Brain Institute (128 tested)

Neurosphere cultures are derived from fetal mouse brain (E13-15) and are grown in defined, serum-free medium

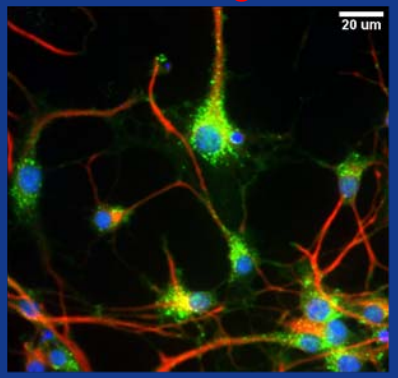


Neuropheres contain CNS stem cells, lineage committed precursor cells, and differentiated cells

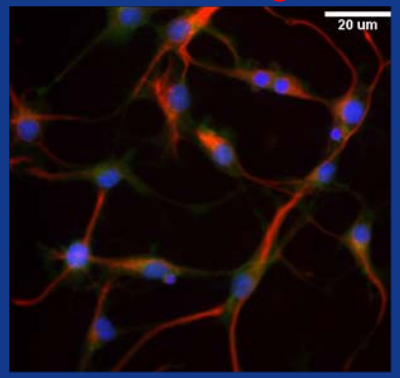


Neurosphere cultures can be infected with RML prions

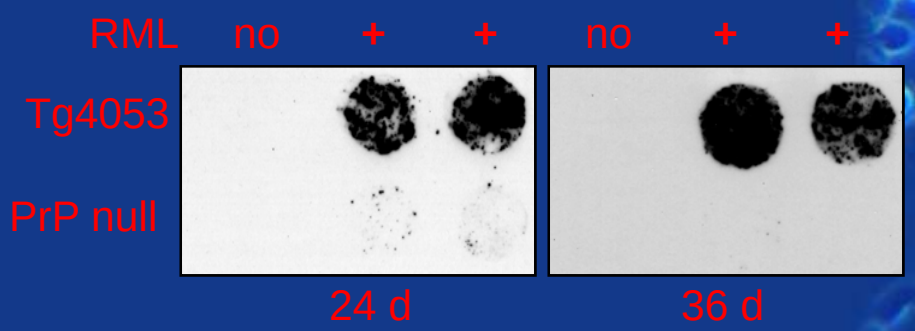
Infected Tg4053



Uninfected Tg4053



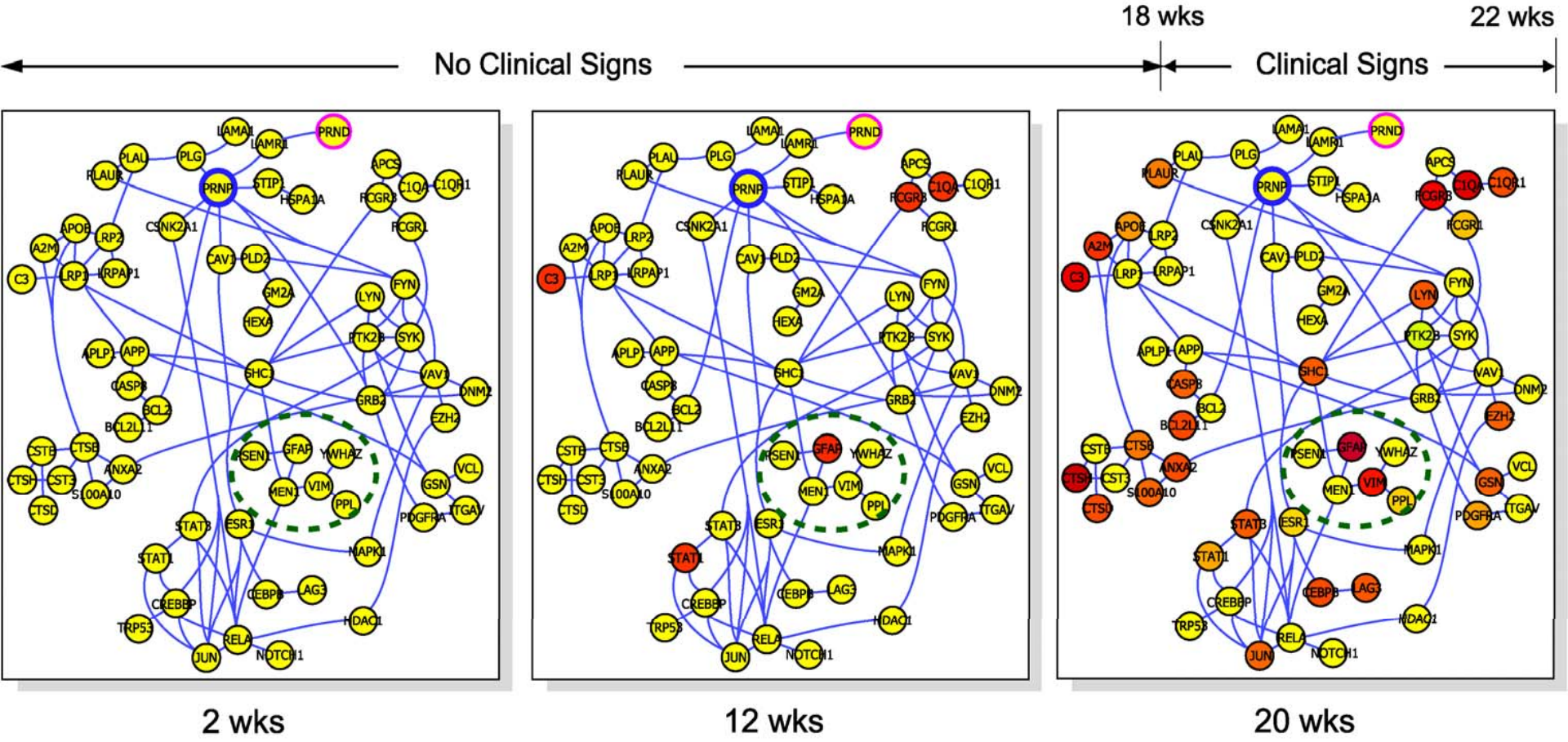
Conformation dependent immunocytochemistry
GdnSCN-treated, nestin, PrP (D18 Fab)



Proteinase K-treated cell blots

Giri et al., 2006, Proc Natl Acad Sci **103**:8875

Dynamics of a Prion Perturbed Network in Mice and Systems Logic for Diagnostics



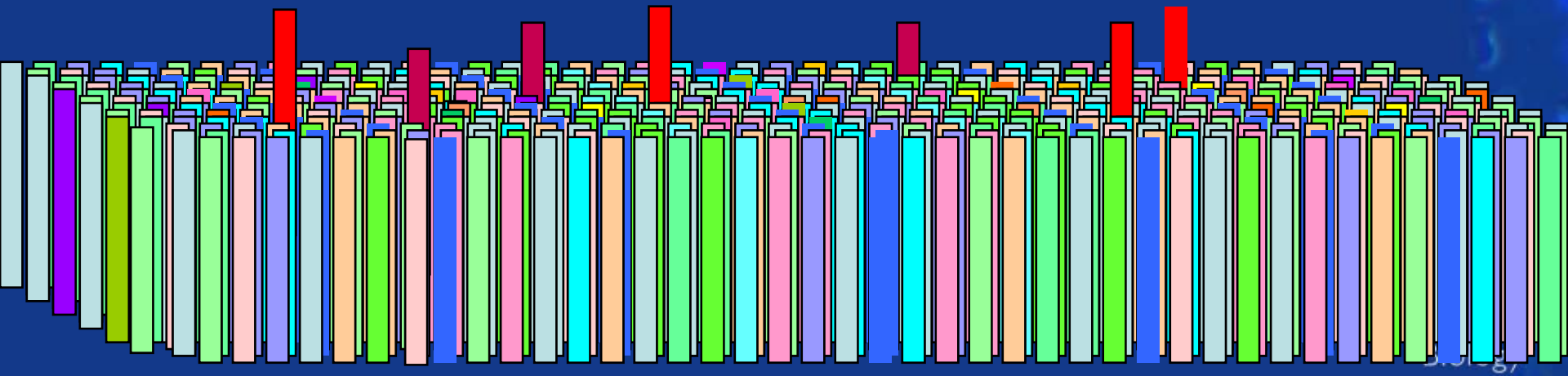
Blood Protein Diagnostics

Finding diagnostic markers:
The 'needle in the haystack' problem

Hundreds of thousands proteins are present (cells, tissue, serum, etc.)

Which one(s) in the serum is/are indicator(s) for a specific disease?

Organ-specific blood markers

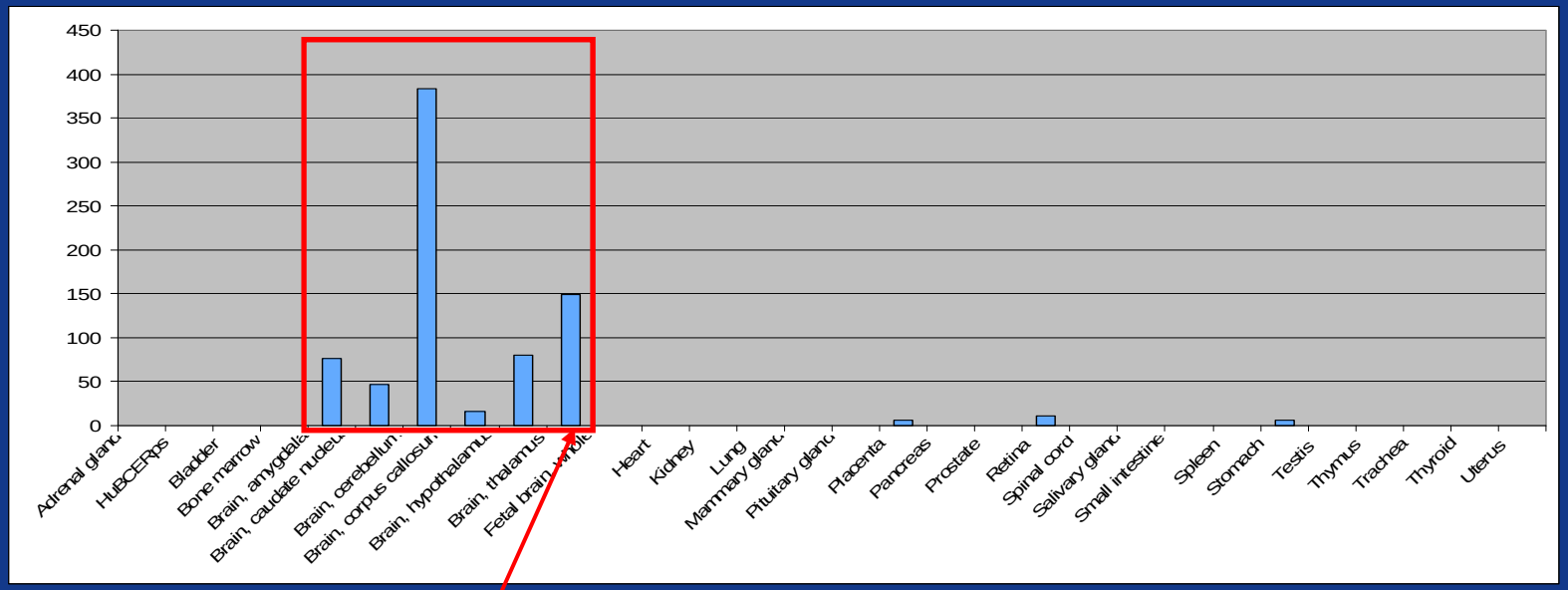


Fingerprinting Prion Disease through the Brain-Specific Blood Proteins

- Out of 900 common DEGs, 100 transcripts are predicted as CNS-specific transcripts by computational searches against transcriptome libraries of MPSS (Massively Parallel Signature Sequencing) on 40 mouse tissues (female).

- 49 CNS-specific gene products are predicted as secreted proteins.

- Some of those marker candidates are being tested by immunoassay methods on serum samples.



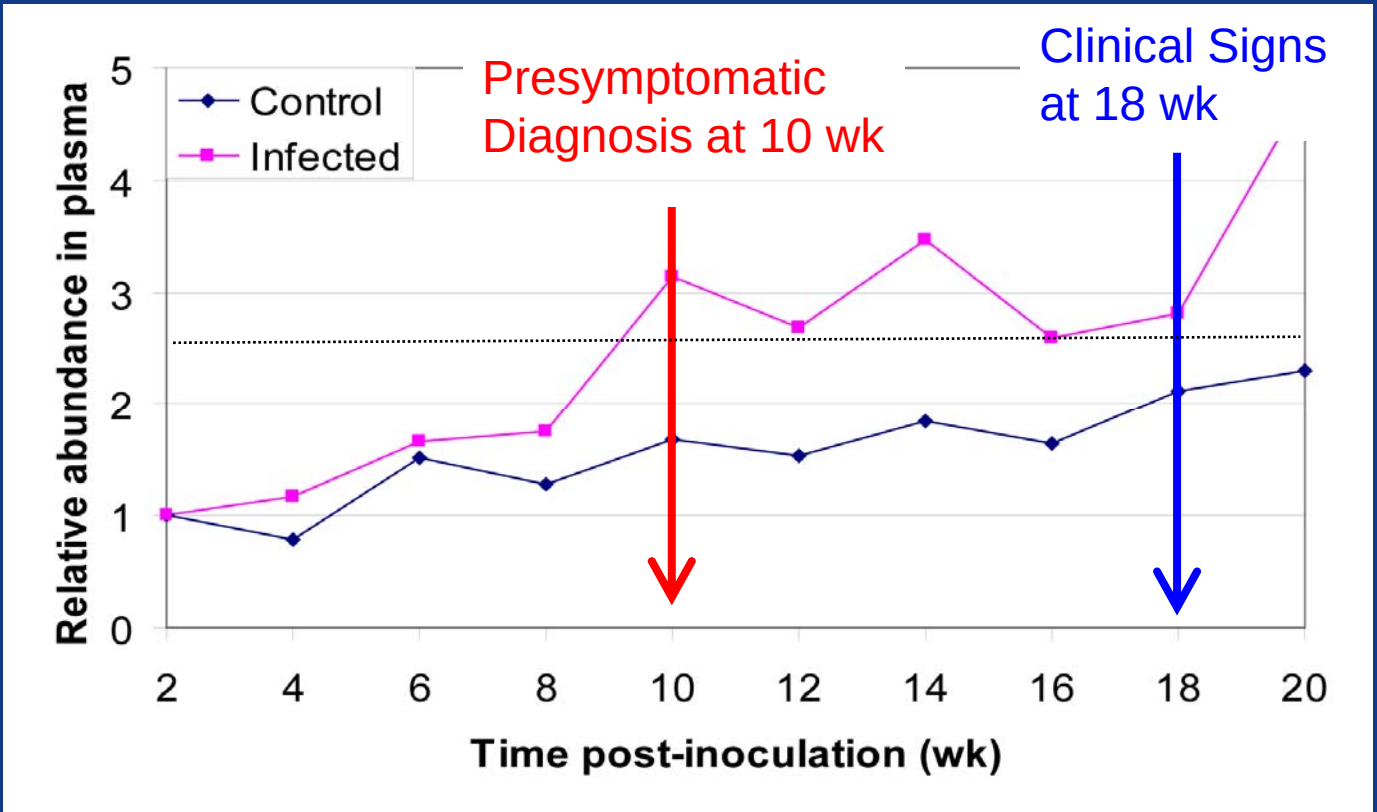
Institute for Systems Biology

Blood Protein Measurements

- Proteomics--mass spectrometry--glycocapture, iTRAC, synthetic peptide standards
- Surface plasmon resonance
- Microfluidics/nanotechnology

Increasing abundance of Protein A from blood samples from prion-infected mice

Dynamic change in relative abundance of “Protein A”

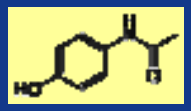


Institute for Systems Biology

Mouse Brain-Specific Blood Proteins

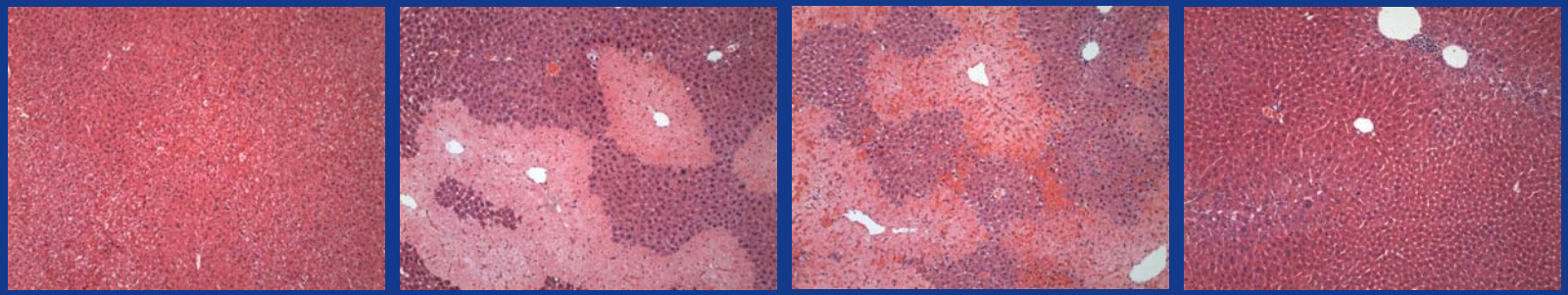
- 4500 mouse blood proteins detected by mass spectrometry
- 2200 blood proteins encoded by transcripts identified by brain MPSS analyses
- 89 of these proteins are encoded by brain-specific blood transcripts
- 87/ mouse blood proteins have human orthologs

Mouse liver toxicity



Acetaminophen Liver Toxicity

Histological confirmation of liver injury and recovery post i.p. APAP



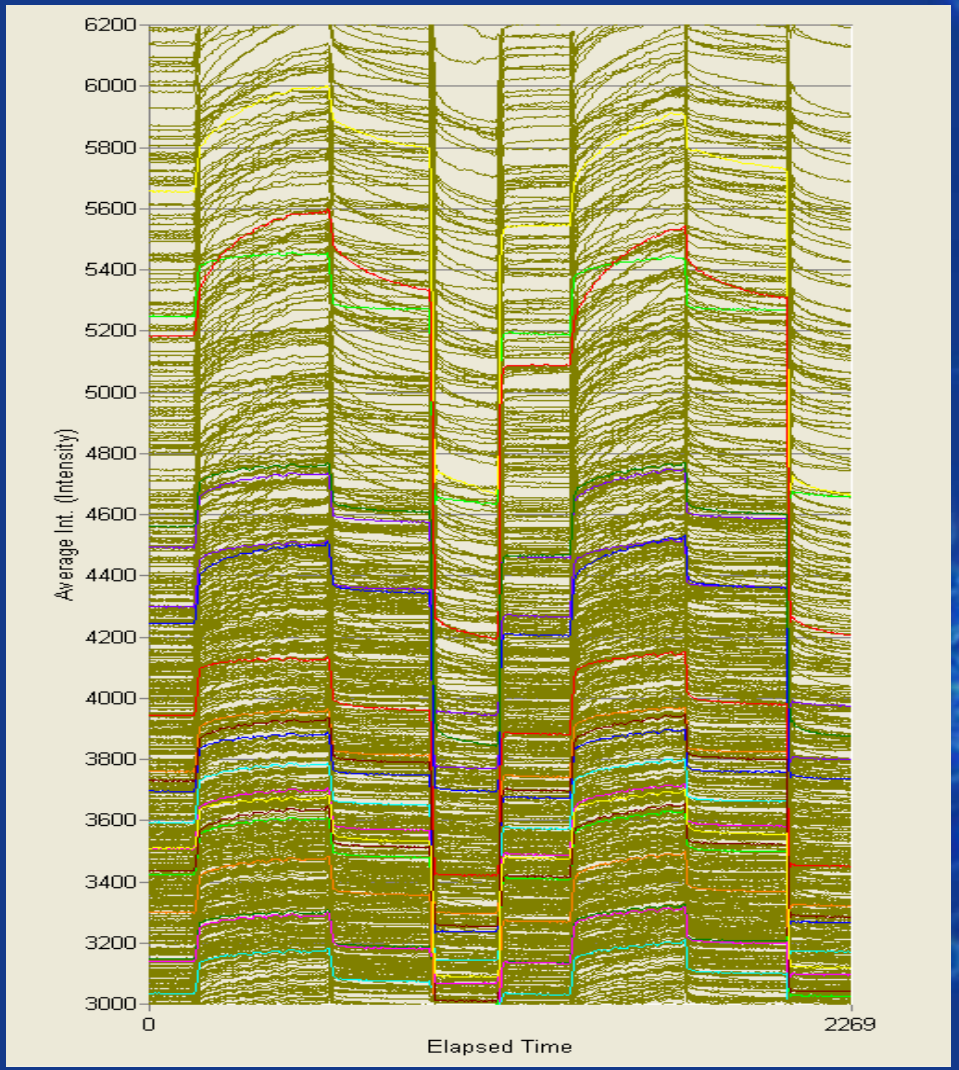
0h 24h 48h 72h

INSTITUTE FOR Systems Biology

INSTITUTE FOR Systems Biology

Raw image of 800 spots array
(400 antibodies in duplicates)

Raw SPR sensorgrams from 800-feature array



Advantages in biomarker discovery

- Label-free---**simple**
- 800 spots---**high throughput**
- 10 minutes for each serum sample---**fast**
- Antibody chip can be regenerated---**low cost**
- 30 samples by same chip---**low variation**

Cross reactivities--multipara-meter analyses

Platform taking few lambda blood samples coming

INSTITUTE FOR Systems Biology

Blood is a window into health and disease

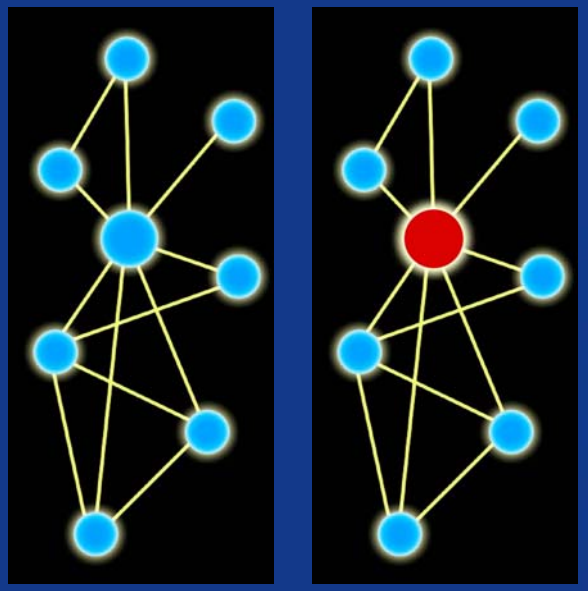
Organ-Specific Blood Proteins Will Make the Blood a Window into Health and Disease

- Perhaps 50 major organs or cell types--each secreting protein blood molecular fingerprint.
- The levels of each protein in a particular blood fingerprint will report the status of that organ. Probably need perhaps 50 organ-specific proteins per organ.
- Will need to quantify 2500 blood proteins from a droplet of blood--use microfluidic/nanotechnology approaches.
- Key point: changes in the levels of organ-specific markers can assess virtually **all** diseases challenges for a particular organ

Blood Is a Window into Health and Disease

- Each human organ has an organ-specific blood molecular fingerprint composed of secreted proteins that reports the state of the organ (normal vs disease).

- The blood molecular fingerprints can be read today (discovery--10s proteins on a few controls) with antibodies and proteomics. Tomorrow they (typing--1000s proteins on millions of patients) will be read with microfluidics and nanotechnology devices.



Non-Diseased Diseased

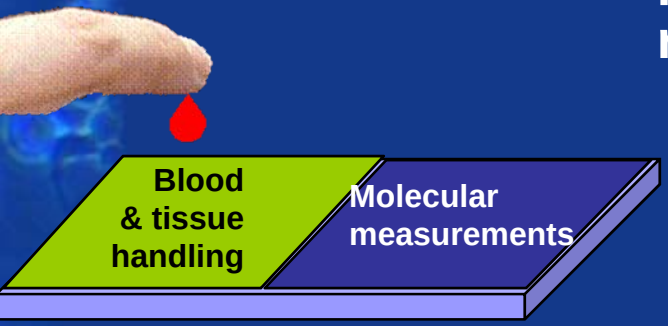
Microfluidic and nanotechnology approaches to blood protein measurements

INSTITUTE FOR Systems Biology

In vitro diagnostics

Quantitate 2500 organ-specific proteins to:

identify disease;
stratify disease;
progression of disease;
response of disease to therapy etc.



Fundamental Materials/Chemical Issues

- Scalable & Simple Detection Technologies
- Multiple Functions Integrated onto Microfluidics Chips
- Protein Capture Agents
- Manufacturability

INSTITUTE FOR Systems Biology

Alliance for NanoSystems Biology: Tools for P4 Medicine

CALTECH

Jim Heath	Chemistry--nanotech
Michael Roukes	Physics--nanotech
Steve Quake*	Applied Physics--microfluidic*

Stanford*

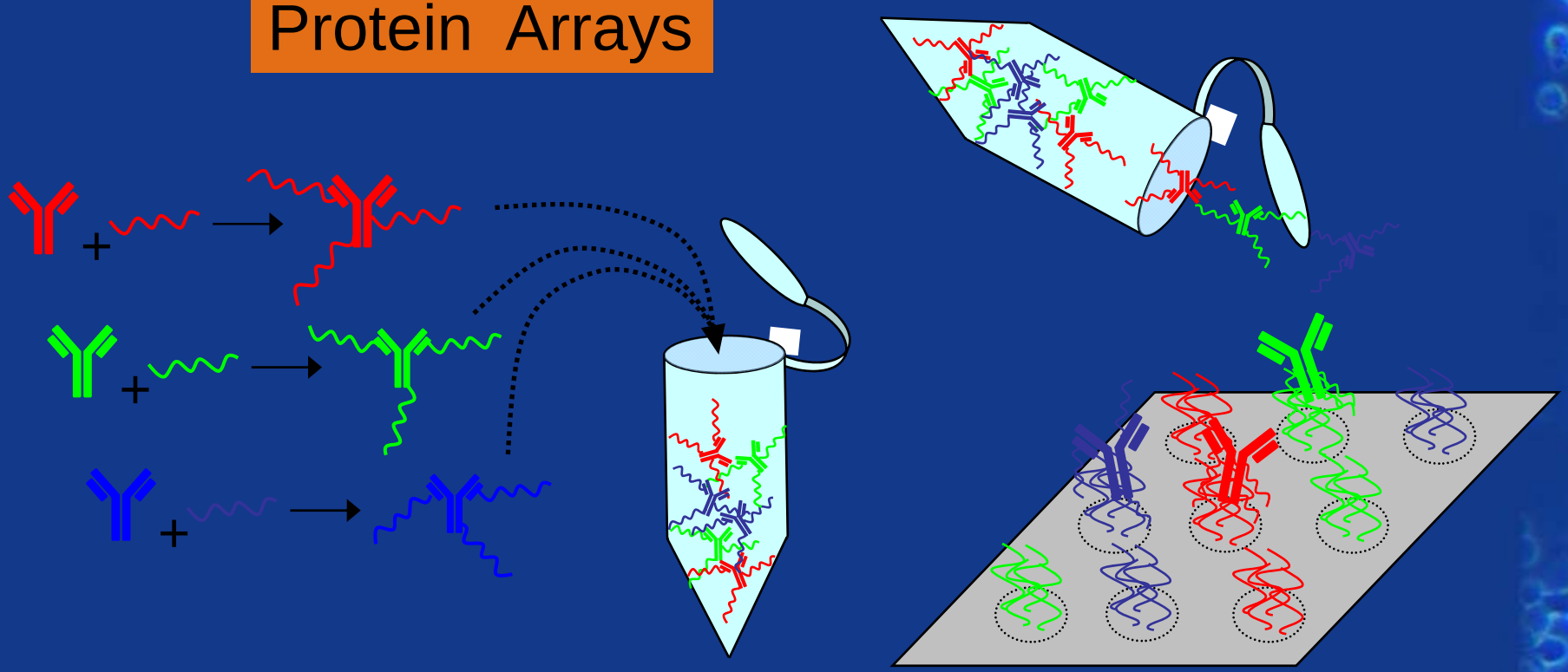
ISB

Lee Hood	Immunology, Technology, Genomics
Alan Aderem	Immunology
Adrian Ozinsky	Microfluidics, Immunology

UCLA

Mike Phelps	Medical Pharmacology--molecular imaging

Protein Arrays



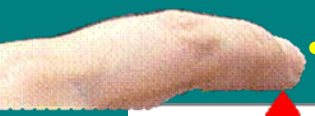
DNA-Encoded Antibody Library scheme. Femptomole sensitivity.

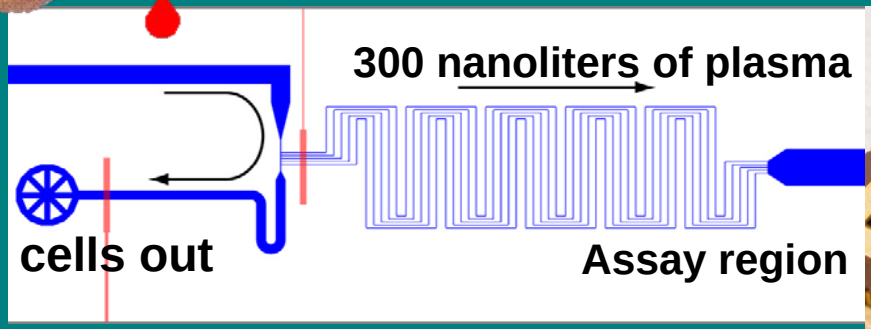
Ryan Bailey and Jim Heath

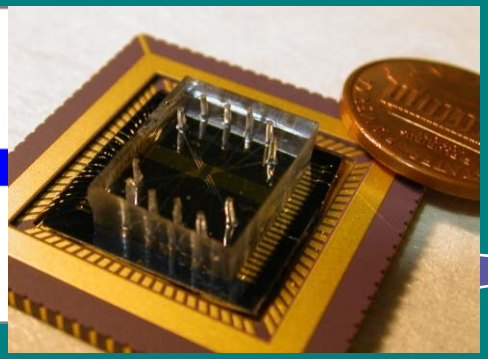
DEAL for *In vitro* molecular diagnostics:
Integrated biology/chemistry/nanotech/microfluidics platforms

Separate plasma & rapidly quantitate protein biomarker panels to:

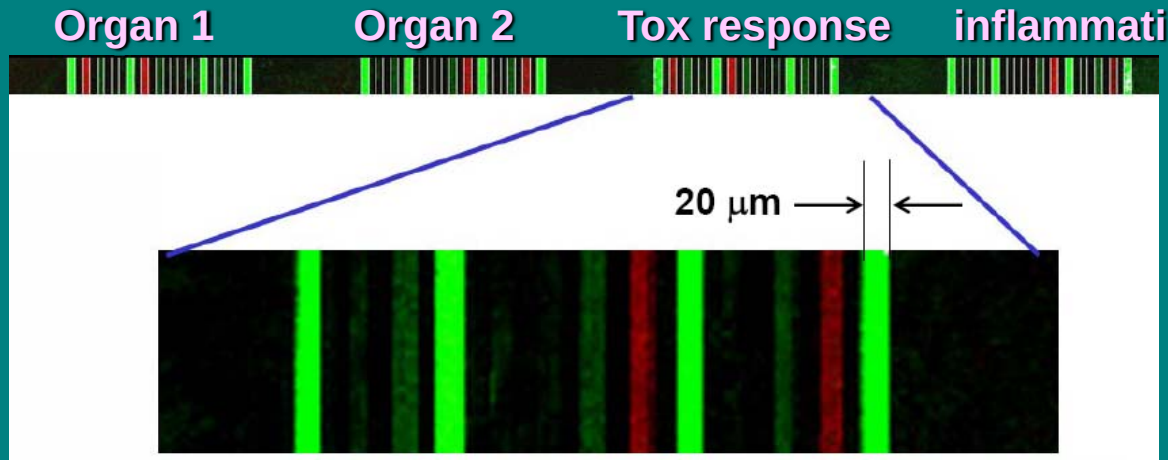
- **Profile health status of individual organs**
- **Detect disease prior to clinical symptoms**
- **Select appropriate therapies or combination therapies**
- **Profile positive & adverse responses to therapies**







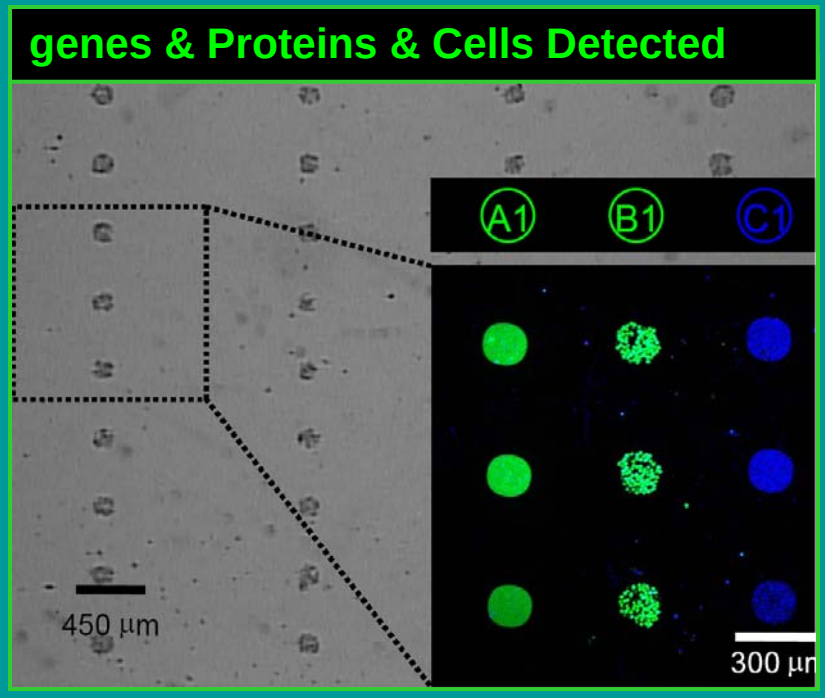
Large panel of protein biomarkers measured in a single microfluidics channel (15 min assay time)



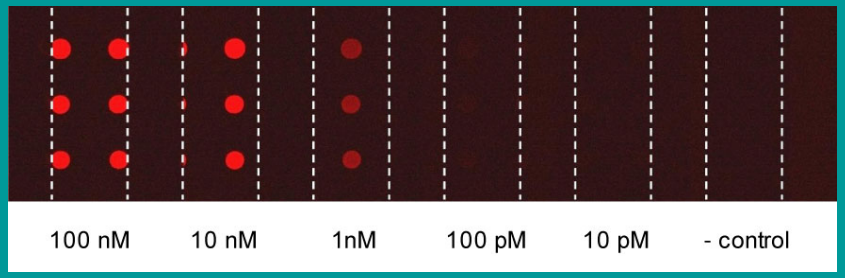
Jim Heath, et al

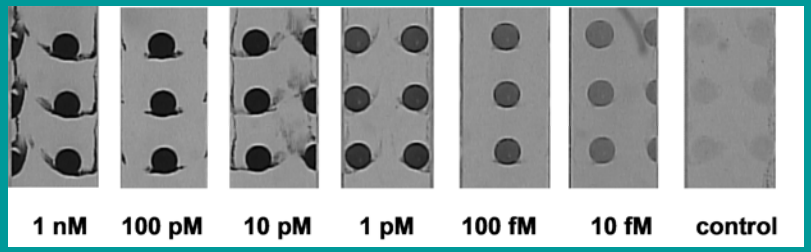
DEAL: DNA-Encoded Antibody Libraries

A single, non-biofouling surface chemistry for gene & protein detection & high efficiency multiplexed cell sorting. Chemistry is thermally stable towards microfluidics fabrication steps (80ºC; few hours) & stable to dehydration



Dynamic range of 10^{-7} – 10^{-15}M





Blood Is a Window into Health and Disease

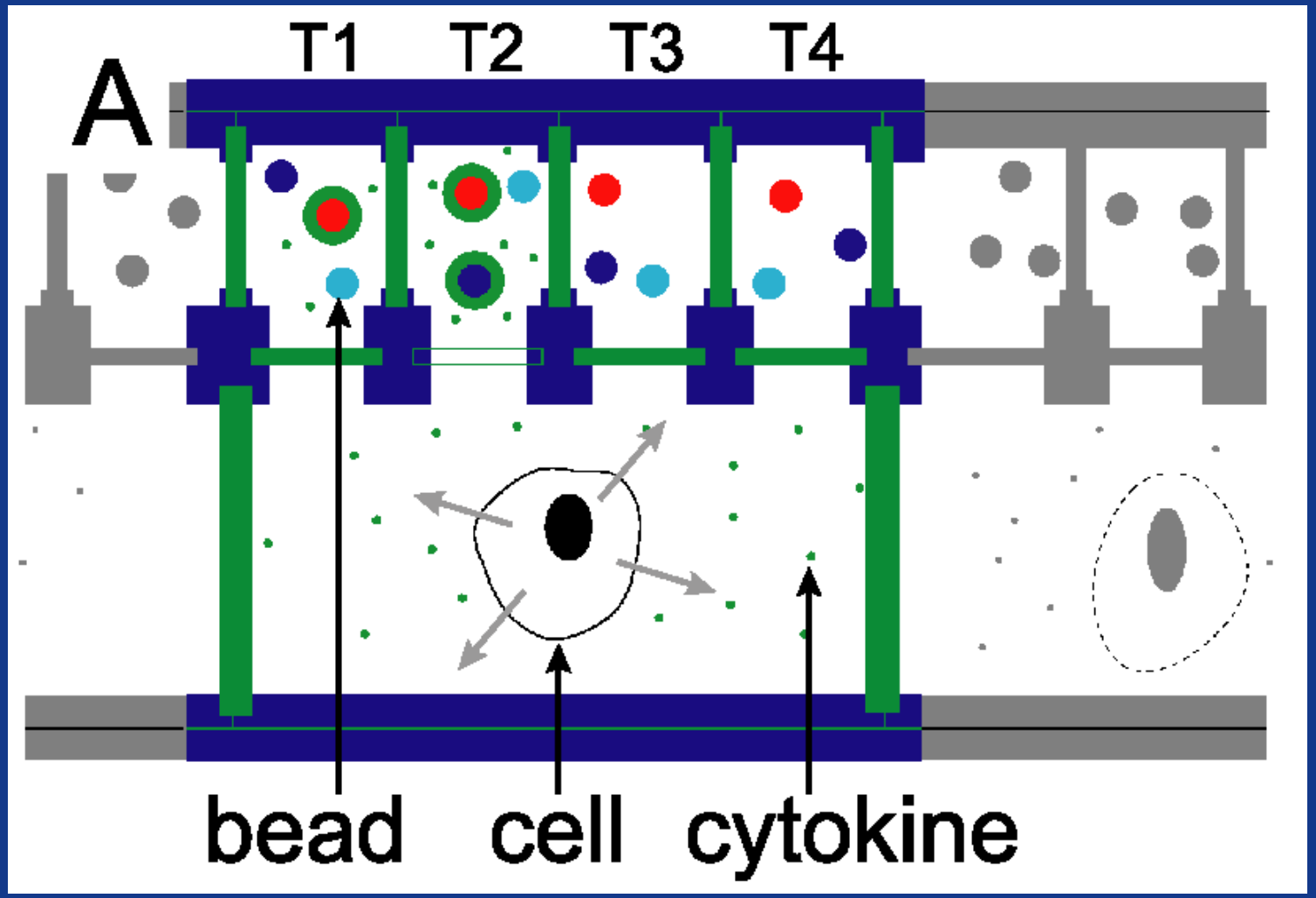
- Read the organ-specific secreted molecular fingerprints
  - Correlation with cellular state change
  - Decipher underlying network changes

- The blood fingerprints will permit:
  - Early diagnosis
  - Disease stratification
  - Follow disease progression
  - Follow response to therapy
  - Early detection of adverse drug reactions--e.g. liver toxicity
  - Check animal disease models for validity vs humans
  - Stratify patients with regard to drug responsiveness
  - Aid in titrating drug doses
  - Monitor drug side reactions
  - Environmental toxin assessment
  - Development
  - Aging

Single Cell Analyses and Single Molecule

Single Cell Analyses

Multiplexed ELISA for secreted proteins



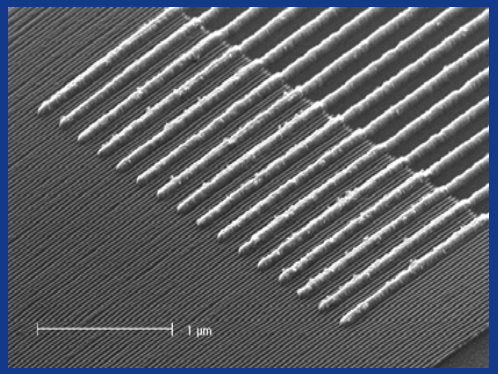
A Nano-Device to Sequence Single DNA Molecules

- Sequence one billion DNA molecules simultaneously from one individual for 100 DNA letters

Personalized Genome Sequences

Predictive, Personalized, Preventive and Participatory (P4) Medicine

- Driven by systems approaches to disease, new measurement (nanotechnology) and visualization technologies and powerful new computational tools, P4 medicine will emerge over the next 10-20 years



Predictive, Personalized, Preventive and Participatory Medicine (P4 Medicine)

- **Predictive:**
  - Probabilistic health history--DNA sequence
  - Biannual multi-parameter blood protein measurements
  - Biannual blood cell measurements--dynamic
  - In vivo and single cell molecular imaging
- **Personalized:**
  - Unique individual human genetic variation mandates individual treatment
  - Patient is his or her own control
  - Perturb blood cells for dynamic measurements
  - Go directly to patient and skip doctor--patient will have all medical information
- **Preventive:**
  - Strategies for re-engineering the behavior of disease-perturbed networks with drugs

    Vaccines

    Focus on wellness
- **Participatory:**
  - Patient understands and participates in medical choices

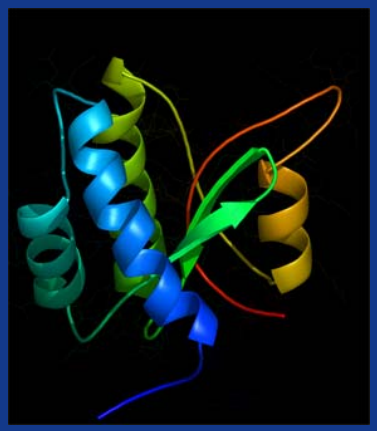
INSTITUTE FOR Systems Biology

Digitalization of Biology and Medicine Will Transform Medicine

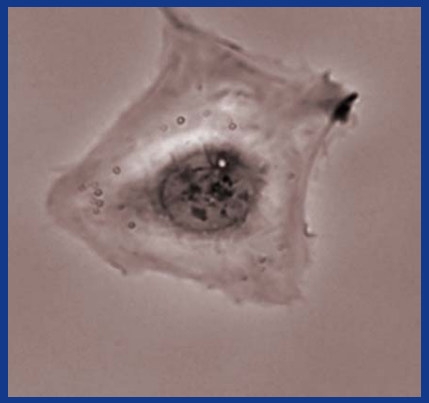
- Analysis of single molecules, single cells and single individuals--*in vitro* and *in vivo*
- A revolution that will transform medicine even more than digitalization transformed information technologies and communications

Single individual

Single molecule



Single cell

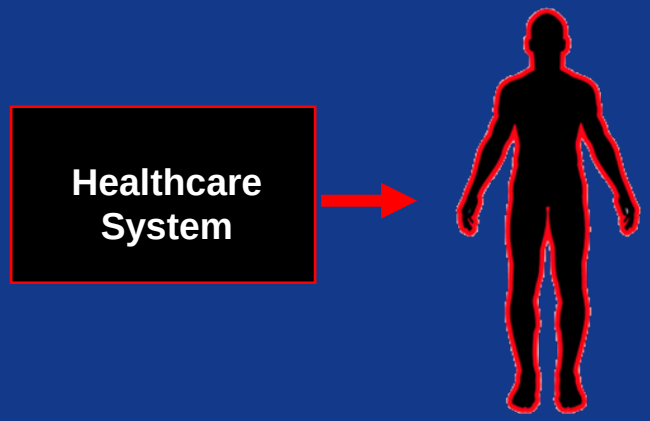


Institute for Systems Biology

P4 Medicine Will Transform the Health Care Industry

Will impact the health care system significantly:

- Pharmaceuticals
- Biotechnology
- IT for healthcare
- Healthcare industry
- Health insurance
- Medicine--diagnostics, therapy, prevention, wellness
- Nutrition
- Assessments of environmental toxicities
- Academia and medical schools



Fundamentally new ideas need
New organizational structures

Institute for Systems Biology

Ten Major Technical Challenges to P4 Medicine

- Methods for individualized genomics – personalized genome sequencing – new methods and approaches.
- Organ-specific proteins in the blood can be transformed into useful biomarkers. Validation of the blood protein molecular fingerprint for each organ.
- Correlate these fingerprints with health and disease states.
- Develop/invent the blood protein measurement technologies
- Imaging molecular disease indications. *In vivo* molecular imaging to follow disease, drug response, drug effectiveness, drug dosage determinations etc.
- Develop new mathematical and computational methods for extracting maximum information from blood fingerprints of individuals, the comparative analyses of genomes from individuals and their molecular images.
- Drug perturbations of networks to be understood in a predictive sense.
- Therapeutic perturbations of biological networks. Re-engineering of networks in higher organisms with drugs (diseased back to normal).
- Handling the enormous personalized data sets—policies, security, mining, reporting, modeling, etc.
- Education of patients, payers, regulators and physicians about P4 medicine.

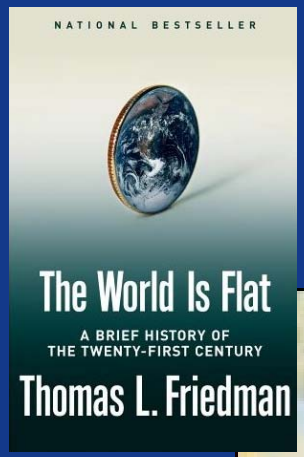
ISB's Concept of a P4 Medical Consortium

- A partnership with PriceWaterhouseCoopers to accelerate and guide the emergence of P4 medicine in the healthcare arena
- Initiate with a drug company, an IT healthcare company, a payer and a provider--to help cast the business plan of the consortium
- Ultimately bring in at least one company from every sector of the health care industry
- ISB-company projects--achieve one or more milestones
- At least one medical school partner willing to embrace systems medicine--and develop and test new approaches
- International ISB-like institutes to bring new resources and support
- Integration of all advances will leverage enormously gains for each consortium member
- Begin critical analysis useful of economic, business, policy and social considerations
- Consortium will raise money to:
  - support key science efforts to attack technical barriers,
  - understand policy, societal, business and economic barriers
  - how much, how fast?

The Flattening of Many Worlds: Globalization of Science

The worlds of science, technology, health are flattening.

Tremendous opportunities for national and international strategic partnerships in science and technology.



- Network of interacting complementary, institutions
  - Training in systems biology and recruiting the best world talent
  - Transferring and collaborating on new technologies and computational tools
  - Strategic partnerships on systems approaches to biology and predictive, prevention, personalized and participatory medicine

INSTITUTE FOR Systems Biology

Some Challenges of Systems Biology and Systems Medicine

- Data space is infinite--carefully formulated perturbations to search relevant data space
- High dimensionality of data--how to reduce high dimensionality to simple hypothesis about biology and disease
- Global measurements of many different data types--assessment
- Capturing data including images from genetic and environmental perturbations to model the dynamics of networks
- Integrating multiscale (hierarchical) types of information
- High dimensional analyses of billions of individual genomes--what biological information can come from the genome alone?
- Granularity of biological information
- There are no pathways, only networks!