Bionature, Lisbon, Portugal
27 March 2013

PANEL BIO:
Are the Current BIO-models Powerful Enough?

Moderator

Son Nghiem, Jet Propulsion Laboratory, California Institute of Technology, USA

Panelists

- **Hiroshi Toyoizumi**, Waseda University, Japan: "Advantage of using applied probability theory in bio-techonolgy"
- **Eduardo dos Santos**, Universidade, Brazil: "Are the current models sufficient to predict the druggable and therapeutic targets that we need?"
- **Hesham Ali**, University of Nebraska, Omaha, USA: "Biological Network"
- **Son Nghiem**, Jet Propulsion Laboratory, California Institute of Technology, USA: "Malaria Model"

Key Points from Panel

- While bio-research can target specific areas or topics, it should be carried out as a part of a coordinated system rather than isolated stand-alone effort in order to optimize the value from the research.
- Bio-modeling should break the traditionally boundary to become extensively interdisciplinary; e.g., to include applied probability theory, computing and internet advances, or satellite observations.
- A system for bio-information should be developed and sustained with contributions from researchers, institutions, and countries with specific protocols for data formats and policies to serve to common interests of academia, governments, and industries.
- Diversified funding supports are important to sustain research system.

Bionature, Lisbon, Portugal
27 March 2013

PANEL BIO: Are the Current BIO-models Powerful Enough?

Moderator

Son Nghiem, Jet Propulsion Laboratory, California Institute of Technology, USA

Panelists

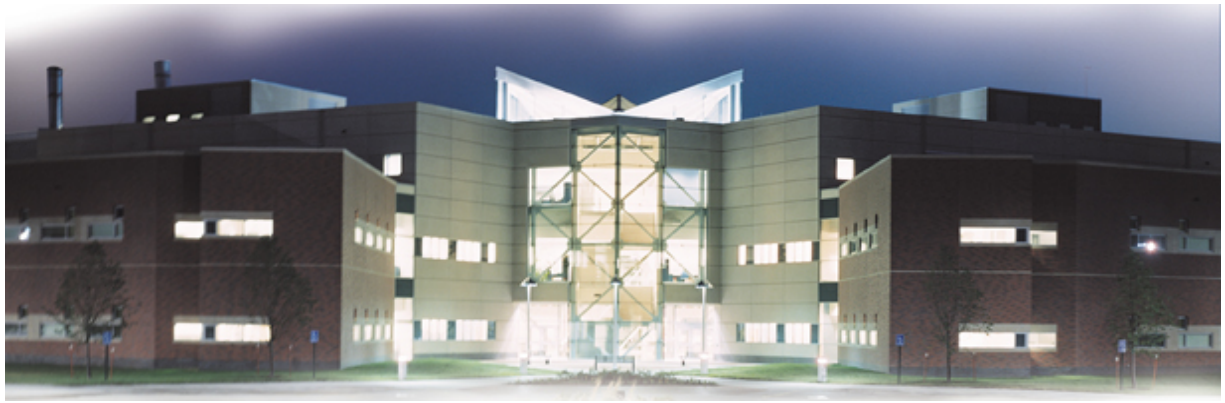
- **Hiroshi Toyoizumi**, Waseda University, Japan: "Advantage of using applied probability theory in bio-techonolgy"
- **Eduardo dos Santos**, Universidade, Brazil: "Are the current models sufficient to predict the druggable and therapeutic targets that we need?"
- **Hesham Ali**, University of Nebraska, Omaha, USA: "Biological Network"
- **Son Nghiem**, Jet Propulsion Laboratory, California Institute of Technology, USA: "Malaria Model"

Key Points from Panel

- While bio-research can target specific areas or topics, it should be carried out as a part of a coordinated system rather than isolated stand-alone effort in order to optimize the value from the research.
- Bio-modeling should break the traditionally boundary to become extensively interdisciplinary; e.g., to include applied probability theory, computing and internet advances, or satellite observations.
- A system for bio-information should be developed and sustained with contributions from researchers, institutions, and countries with specific protocols for data formats and policies to serve to common interests of academia, governments, and industries.
- Diversified funding supports are important to sustain research system.

University of Nebraska at Omaha

Are the Current Bio-models Powerful Enough?



BIOTECHNO 2013

Hesham H. Ali

UNO Bioinformatics Core Facility

College of Information Science and Technology

Biomedical Informatics – Where are we?

- Availability of many large useful database systems; private and public
- Advances in new technologies as high throughput next generation sequencing
- Availability of numerous helpful software tools
- Fragmented, in some case isolated, efforts by computational scientists and bioscientists
- The trendiness of the discipline
- Massive interest from Industry, researchers and the public
- Many advances in Biomedical research

Simple Questions

Where is the cure for cancer?

Why don't we have personalized medicine?

Why is AIDS still misunderstood?

Can effectively be boiled down to:

Why hasn't high-throughput data been effectively harnessed yet?

Plausible Answers

It is not that easy:

- Complexity of the system
- Complexity of the organisms
- Size of the data (“big data”)
- Search space of inter-data relationships
- Heterogeneity of the data
- Lack of integration of data
- Computing power
- …..

Robin Roberts: I will beat MDS Updated Mon June 11, 2012

"Good Morning America's" Robin Roberts is bravely facing a new health battle. The 51-year-old revealed Monday that five years after overcoming breast **cancer**, she's been diagnosed with a rare blood disorder that affects the bone marrow called...

http://marquee.blogs.cnn.com/2012/06/11/robin-roberts-i-will-beat-mds/

Comedian Tommy Chong fighting prostate **cancer** Updated Sun June 10, 2012

Tommy Chong, one-half of the marijuana-loving "Cheech and Chong" comedy duo, is battling prostate **cancer**, he announced Saturday on CNN.

http://www.cnn.com/2012/06/09/showbiz/chong-prostate-cancer/index.html

What can be done about the deepening polarization in America? Updated Wed June 6, 2012

By CNN's Jack Cafferty: The polarization of America is like a **cancer** that is slowly killing us. And like many forms of **cancer**, there appears to be no **cure**. We are more severely divided now than at any time in the last 25 years according to a new pew...

http://caffertyfile.blogs.cnn.com/2012/06/06/what-can-be-done-about-the-deepening-polarization-in-america/

Experimental drug offers new way to battle certain breast **cancer** Updated Sun June 3, 2012

Doctors who treat breast **cancer** patients are very excited about an experimental drug that presents a whole new way of knocking out **cancer** cells.

http://www.cnn.com/2012/06/03/health/breast-cancer-drug/index.html

But….

- Her2+ BC
  - 20-25% of breast cancers
  - Normal treatment: Herceptin
    - Eventually stops working if cancer comes back
- T-DM1 Drug
  - Trojan horse drug
  - 3 extra months of improvement
  - Lack of usual side effects

UNO BIOINFORMATICS

Personalized Medicine

Resource

Cell

Personal Omics Profiling Reveals Dynamic Molecular and Medical Phenotypes

Rui Chen,[1,11] George I. Mias,[1,11] Jennifer Li-Pook-Than,[1,11] Lihua Jiang,[1,11] Hugo Y.K. Lam,[1,12] Rong Chen,[2,12] Elana Miriami,[1] Konrad J. Karczewski,[1] Manoj Hariharan,[1] Frederick E. Dewey,[3] Yong Cheng,[1] Michael J. Clark,[1] Hogune Im,[1] Lukas Habegger,[6,7] Suganthi Balasubramanian,[6,7] Maeve O'Huallachain,[1] Joel T. Dudley,[2] Sara Hillenmeyer,[1] Rajini Haraksingh,[1] Donald Sharon,[1] Ghia Euskirchen,[1] Phil Lacroute,[1] Keith Bettinger,[1] Alan P. Boyle,[1] Maya Kasowski,[1] Fabian Grubert,[1] Scott Seki,[2] Marco Garcia,[2] Michelle Whirl-Carrillo,[1] Mercedes Gallardo,[9,10] Maria A. Blasco,[9] Peter L. Greenberg,[4] Phyllis Snyder,[1] Teri E. Klein,[1] Russ B. Altman,[1,5] Atul J. Butte,[2] Euan A. Ashley,[3] Mark Gerstein,[6,7,8] Kari C. Nadeau,[2] Hua Tang,[1] and Michael Snyder[1,*]

[1]Department of Genetics, Stanford University School of Medicine

[2]Division of Systems Medicine and Division of Immunology and Allergy, Department of Pediatrics

[3]Center for Inherited Cardiovascular Disease, Division of Cardiovascular Medicine

[4]Division of Hematology, Department of Medicine

[5]Department of Bioengineering

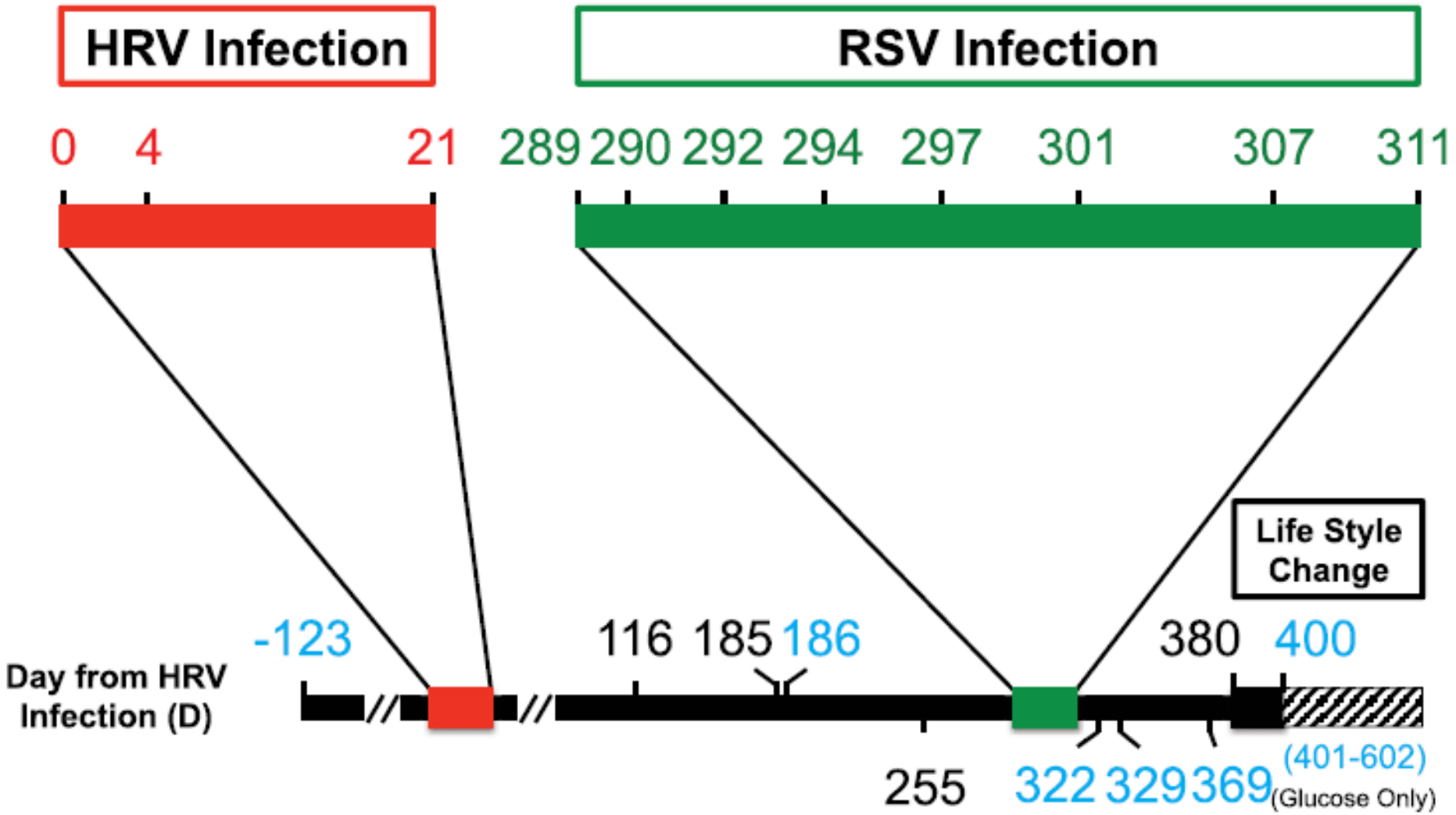
Stanford University, Stanford, CA 94305, USA

Methods

- 54yr old male volunteer
- Plasma and serum used for testing
- 14 month time course
- Complete medical exams and labs at each meeting (20 time points total)
- Extensive sampling at 2 periods of viral infection:
  - HRV (human rhinovirus) – common cold
  - RSV (respiratory synticial) - bronchitis

Time-course summary

726 days total
HRV - red
RSV - green
Fasting - blue
Lifestyle change: ↑exercise, took ibuprofen daily, ↓sugar intake



Deep Dynamic Omics Analysis

- Transcriptome: RNA-Seq of 20 time points
  - 2.67 billion paired end reads
  - 19,714 isoforms for 12,659 genes tracked
- Proteome: Quantification of 6,280 proteins
  - 14 time points via TMT and LC/MS
- Metabolome: 1,020 metabolites tracked during viral infections
  - miRNA analysis also during HRV infection

Techniques Used

- Summary of techniques used:
  - Sample collection
  - HRV and RSV detection
  - Whole-genome sequencing
  - Whole-exome sequencing
  - Sanger-DNA sequencing
  - Whole-transcriptome sequencing: mRNA-Seq
  - Small RNA sequencing: microRNA-Seq
  - Serum Shotgun Proteome Profiling
  - Serum Metabolome Profiling
  - Serum Cytokine Profiling
  - Autoantibodyome Profiling
  - Telomere Length Assay
  - Genome Phasing
  - Omics Data Analysis

Team Count: 41

Resource

Cell

Personal Omics Profiling Reveals Dynamic Molecular and Medical Phenotypes

Rui Chen,[1,11] George I. Mias,[1,11] Jennifer Li-Pook-Than,[1,11] Lihua Jiang,[1,11] Hugo Y.K. Lam,[1,12] Rong Chen,[2,12] Elana Miriami,[1] Konrad J. Karczewski,[1] Manoj Hariharan,[1] Frederick E. Dewey,[3] Yong Cheng,[1] Michael J. Clark,[1] Hogune Im,[1] Lukas Habegger,[6,7] Suganthi Balasubramanian,[6,7] Maeve O'Huallachain,[1] Joel T. Dudley,[2] Sara Hillenmeyer,[1] Rajini Haraksingh,[1] Donald Sharon,[1] Ghia Euskirchen,[1] Phil Lacroute,[1] Keith Bettinger,[1] Alan P. Boyle,[1] Maya Kasowski,[1] Fabian Grubert,[1] Scott Seki,[2] Marco Garcia,[2] Michelle Whirl-Carrillo,[1] Mercedes Gallardo,[9,10] Maria A. Blasco,[9] Peter L. Greenberg,[4] Phyllis Snyder,[1] Teri E. Klein,[1] Russ B. Altman,[1,5] Atul J. Butte,[2] Euan A. Ashley,[3] Mark Gerstein,[6,7,8] Kari C. Nadeau,[2] Hua Tang,[1] and Michael Snyder[1,*]

[1]Department of Genetics, Stanford University School of Medicine

[2]Division of Systems Medicine and Division of Immunology and Allergy, Department of Pediatrics

[3]Center for Inherited Cardiovascular Disease, Division of Cardiovascular Medicine

[4]Division of Hematology, Department of Medicine

[5]Department of Bioengineering

Stanford University, Stanford, CA 94305, USA

Data Generation vs. Integration/Analysis

- New technologies lead to new data:
  - Competition to have the latest technology
  - Focus on storage needs to store yet more data
- Bioinformatics community needs to move from a total focus on data generation to a blended focus of measured data generation (to take advantage of new technologies) and data analysis/interpretation/visualization
- How do we leverage data? Integratable? Scalable?
- From Data to Information to Knowledge to Decision making

“Big Data”

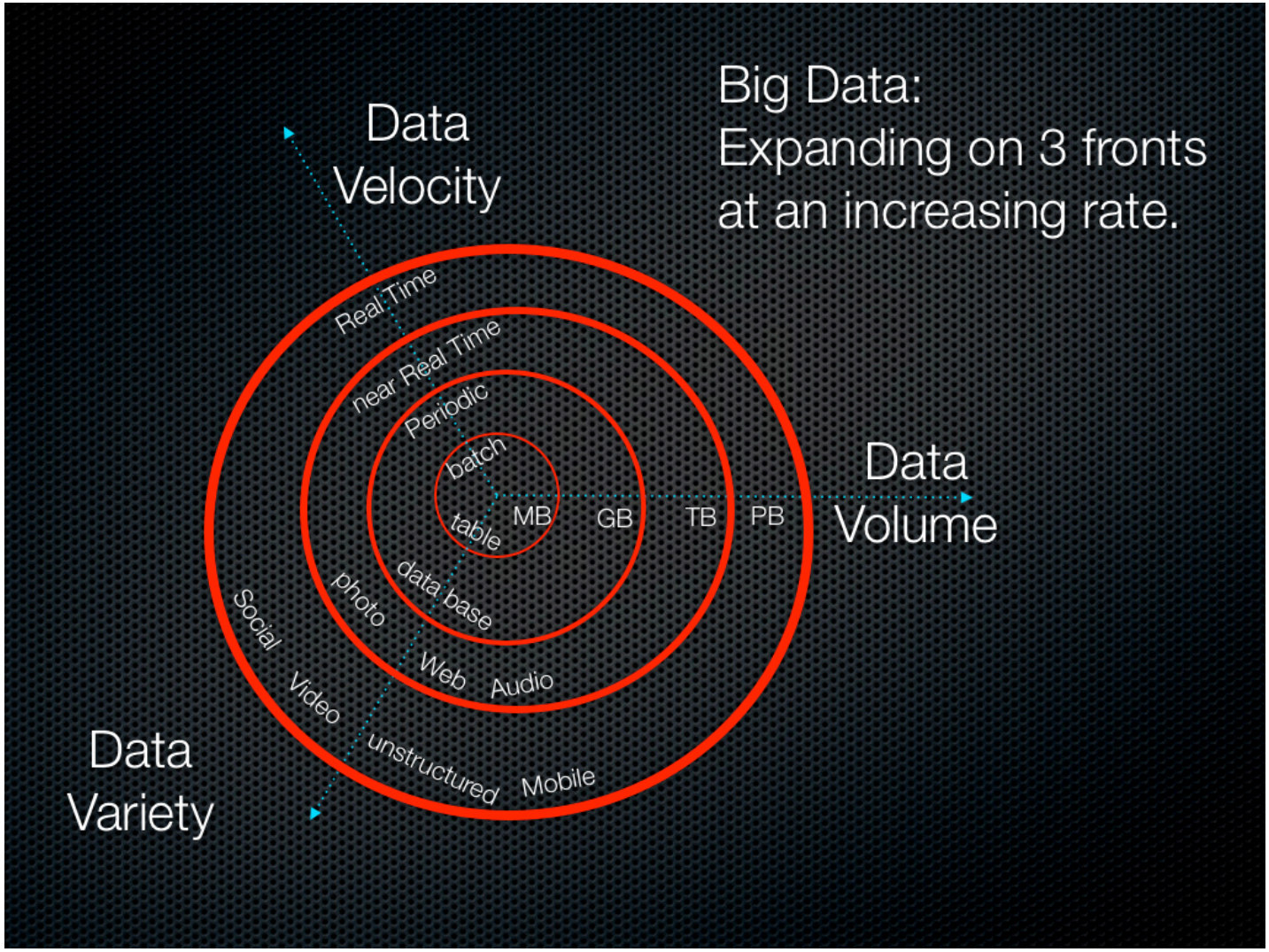


Image Source: http://api.ning.com/files/O6-JOcfS6sxRuZ8I2i5nJVa59xL-krT-a6UqeoLNaHwL2w-JSR-Cy56PmikOvwROgv2gDfYxLAb0Hs*VFr8IePv5OFBJdhDH/BigData.001.jpg

Data versus Knowledge

- With high throughput data collection, Biology needs ways not only to store data but also to store knowledge (Smart data)
- Data: Things that are measured
- Information: Processed data
- Knowledge: Processed data plus meaningful relationships between measured entities

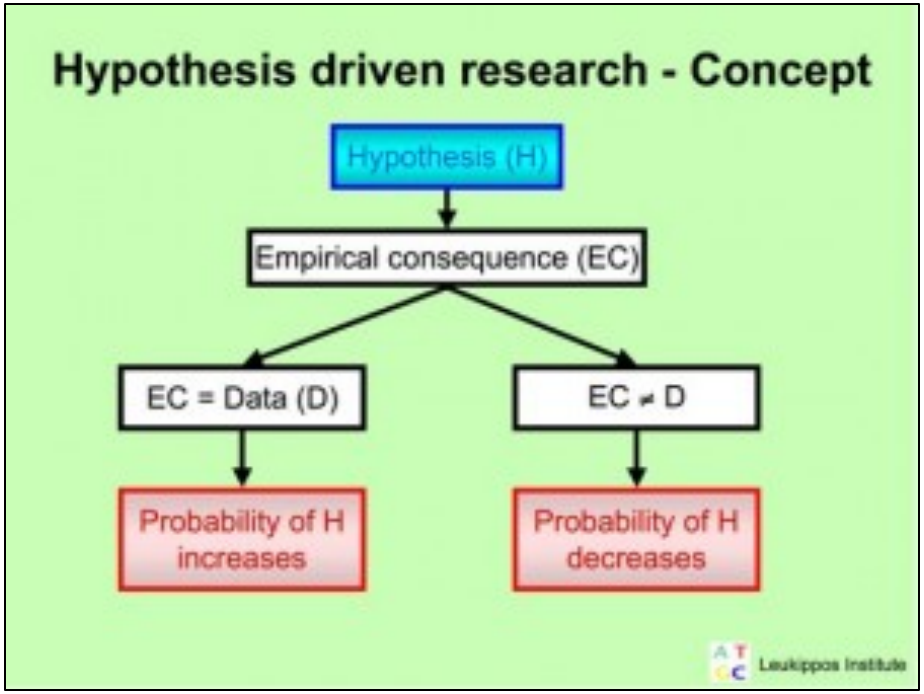
  Power of graph modeling

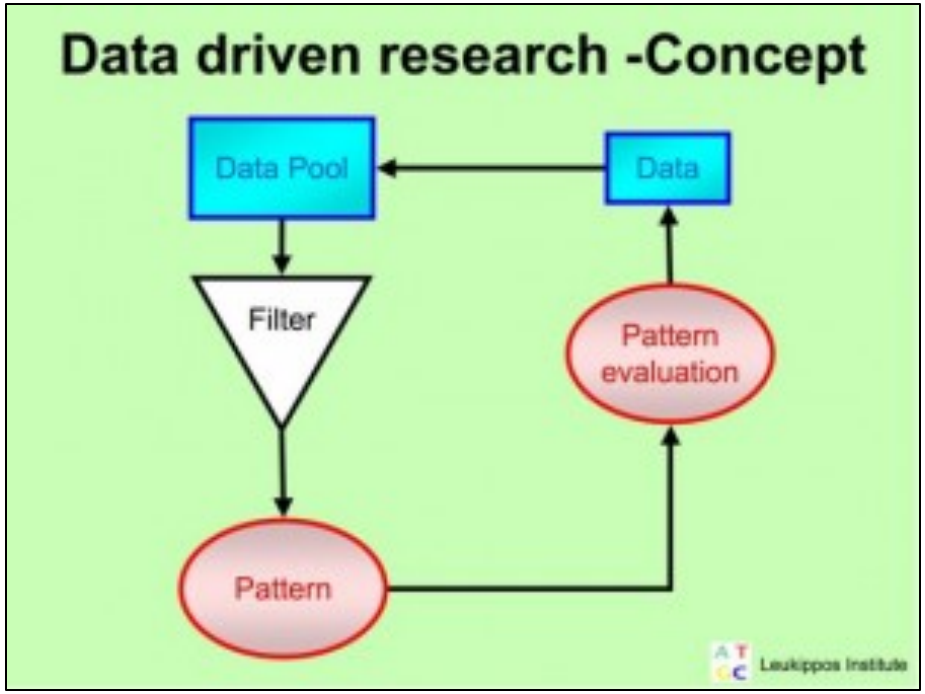
Data-Driven Decisions

- With high throughput data collection, Biology needs ways not only to store data but also to store knowledge (Smart data)
- Data: Things that are measured
- Information: Processed data
- Knowledge: Processed data plus meaningful relationships between measured entities
- Decision Support

A Potential Major Change

- Data driven research vs. Hypothesis driven research





http://www.genome-engineering.com/from-hypothesis-to-data-driven-research.html

Systems Biology Approach

UNO BIOINFORMATICS

- Realistic and Innovative:
  - Networks model relationships, not just elements
  - Discover groups of relationships between genes and gene products

- Validation and Discovery Aspects
  - Examine changes in systems
    - Normal vs. diseased
    - Young vs. old
    - Stage I v. State II v. Stage III v. Stave IV

Systems Biology

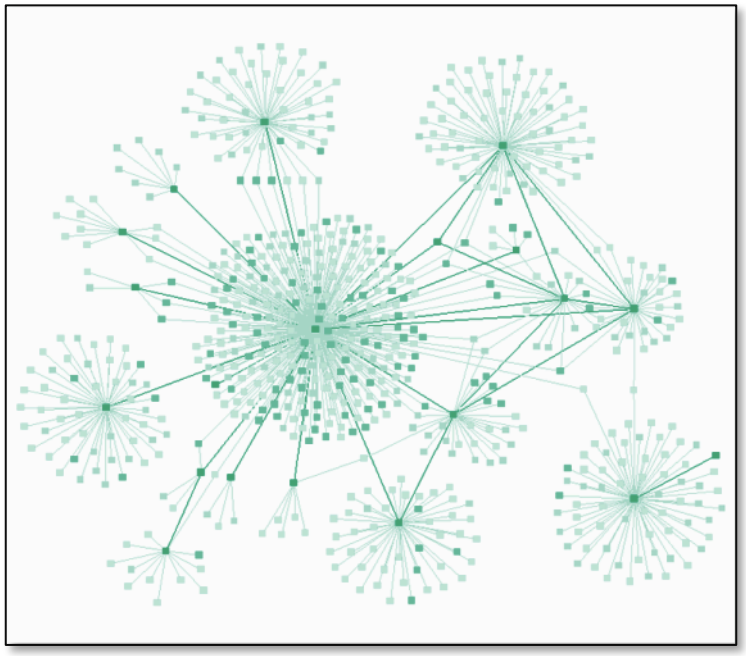
- Holist view of the system
  - Ability to zoom in/out to view critical system components
- Past: Reductionist biology
  - Find a gene/protein of interest
  - Examine under different conditions
- Systems biology: examine an entire system at different conditions

Biological Networks

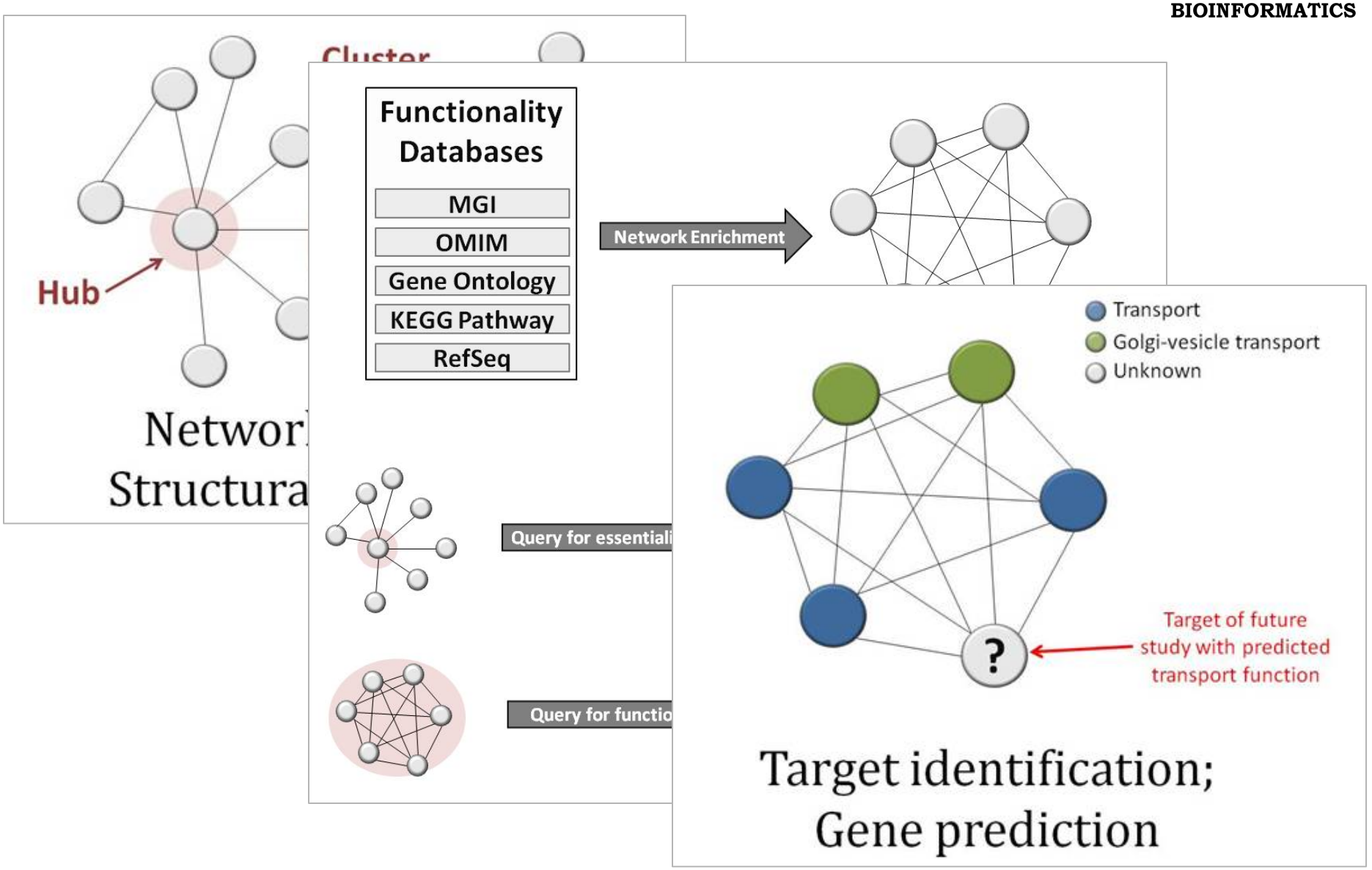
- A biological network represents elements and their interactions

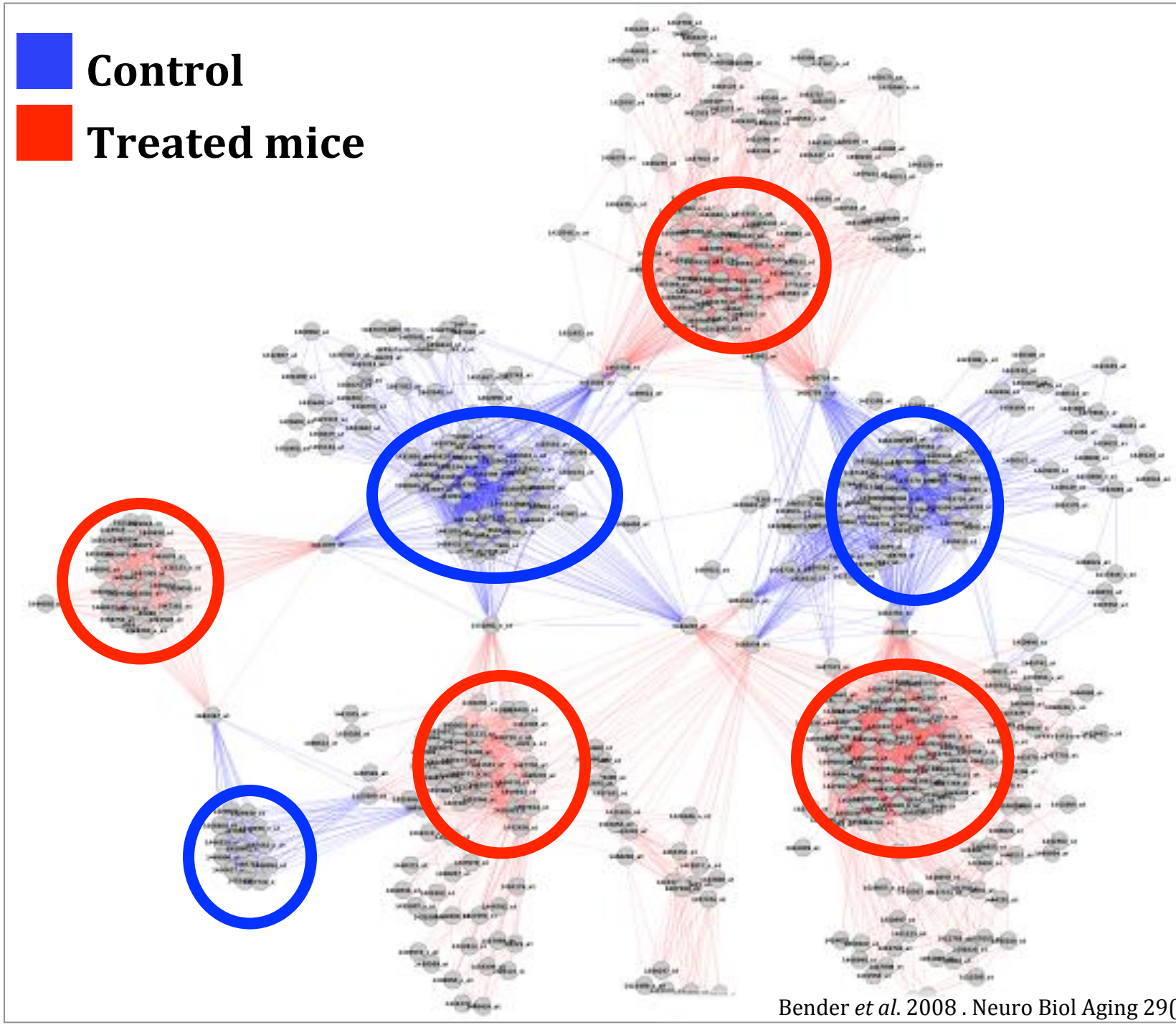
- Nodes → elements
- Edges → interactions

- Can represent multiple types of elements and interactions

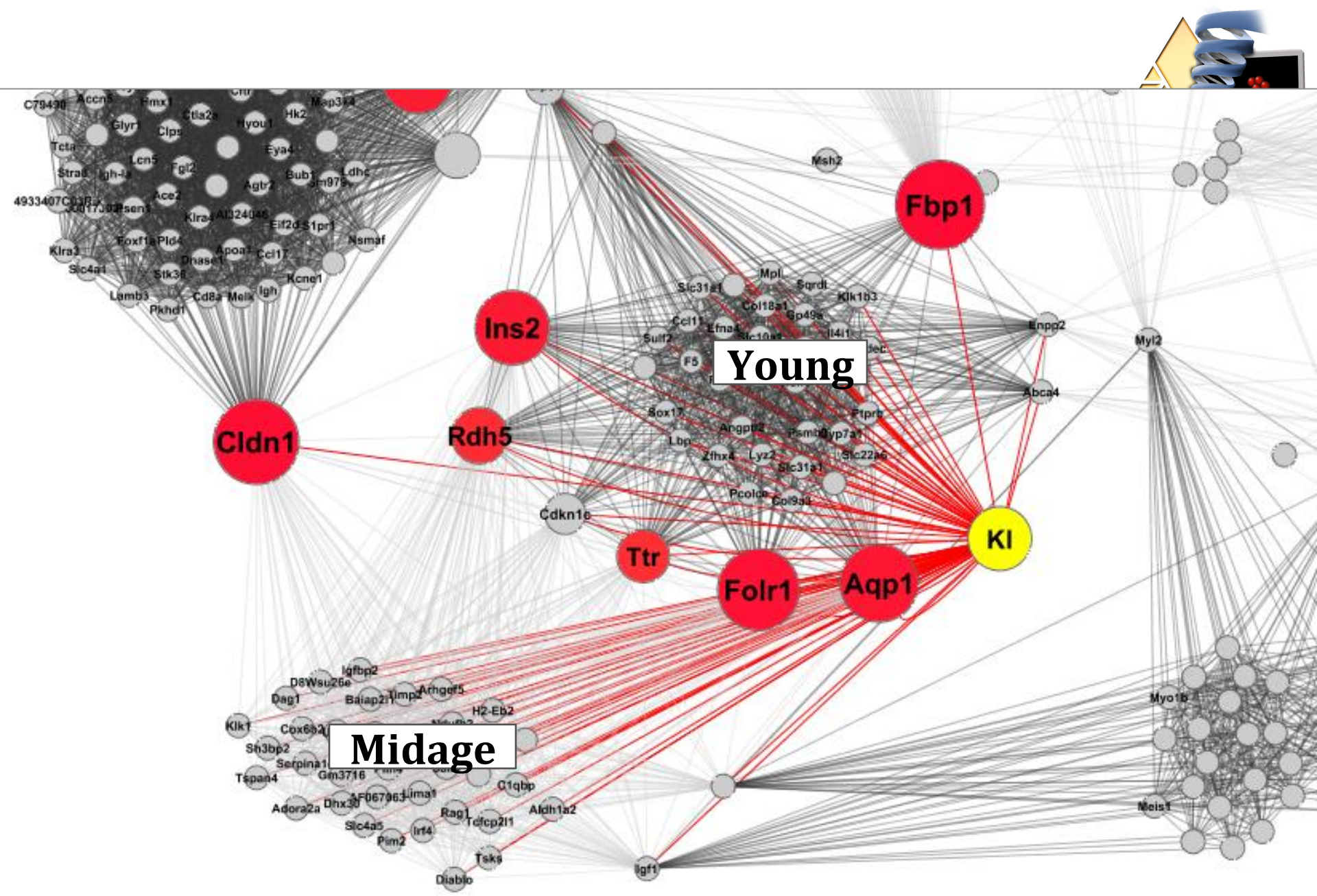


Integrated Data Model

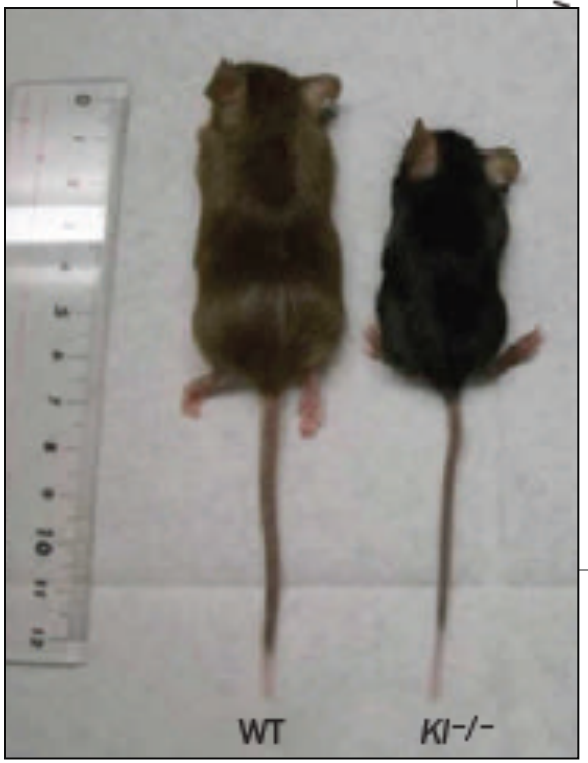


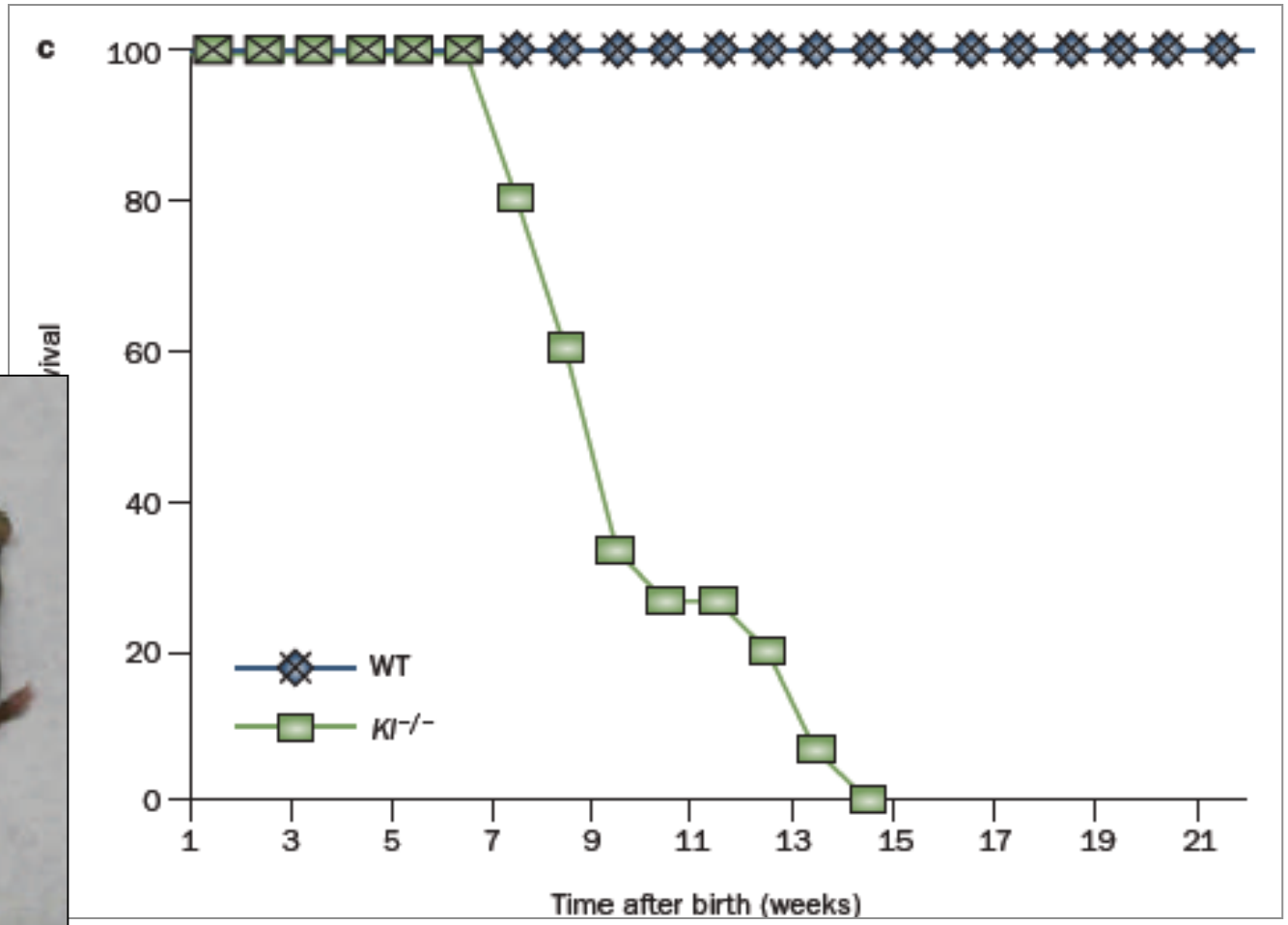


Bender *et al.* 2008 . Neuro Biol Aging 29(9):1404-11


Young
Midage
Fbp1
Ins2
Cldn1
Rdh5
Ttr
Folr1
Aqp1
Kl
Cdkn1c
Enpp2
Abca4
Myl2
Myo1b
Meis1
Igf1
Msh2
Slc31a1
Mpl
Sqrdl
Klk1b3
Col18a1
Gp49a
Ccl11
Efna4
Il4i1
Sulf2
F5
Sox17
Ptprb
Lbp
Zfhx4
Lyz2
Slc22a6
Pcolce
Col9a3
Igfbp2
D8Wsu26e
Dag1
Baiap2l1
Timp2
Arhgef5
H2-Eb2
Klk1
Cox6b2
Sh3bp2
Serpina1
Tspan4
Gm3716
C1qbp
Adora2a
Dhx30
Rag1
Tcfcp2l1
Aldh1a2
Slc4a5
Irf4
Pim2
Tsks
Diablo
Accn5
Fmx1
Ctla2a
Hyou1
Hk2
Map3k4
Glyr1
Clps
Tcta
Eya4
Lcn5
Fgl2
Bub1
Ldhc
Agtr2
Ace2
Klra4
Eif2d
S1pr1
Foxf1a
Pld4
Nsmaf
Klra3
Apoa2
Ccl17
Dnase1
Slc4a1
Stk38
Kcne1
Lamb3
Cd8a
Melk
Igh
Pkhd1

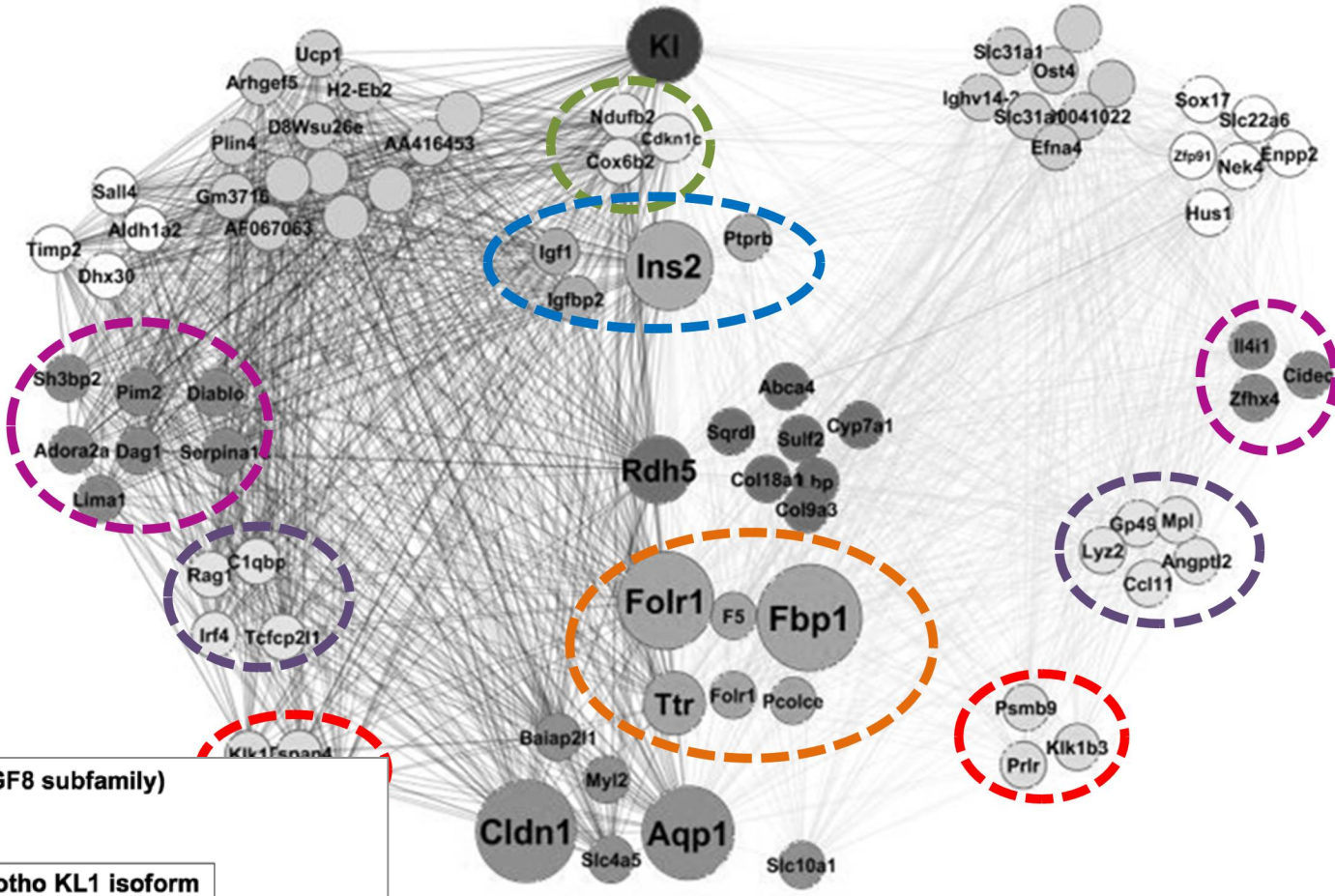
High BD Node: Validation

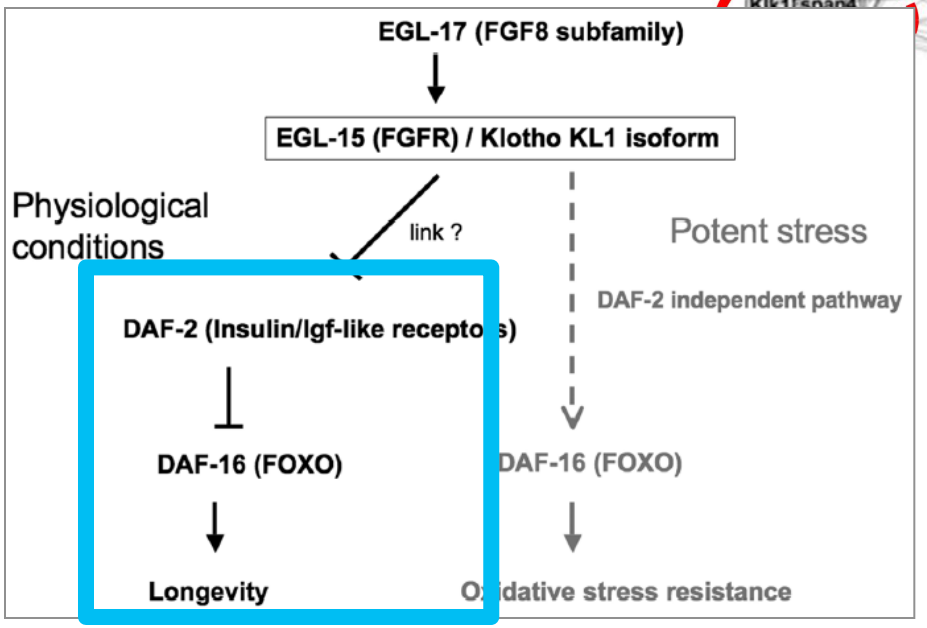




Razzaque,2009. Nat. Rev. Endocrin. 5:611-619

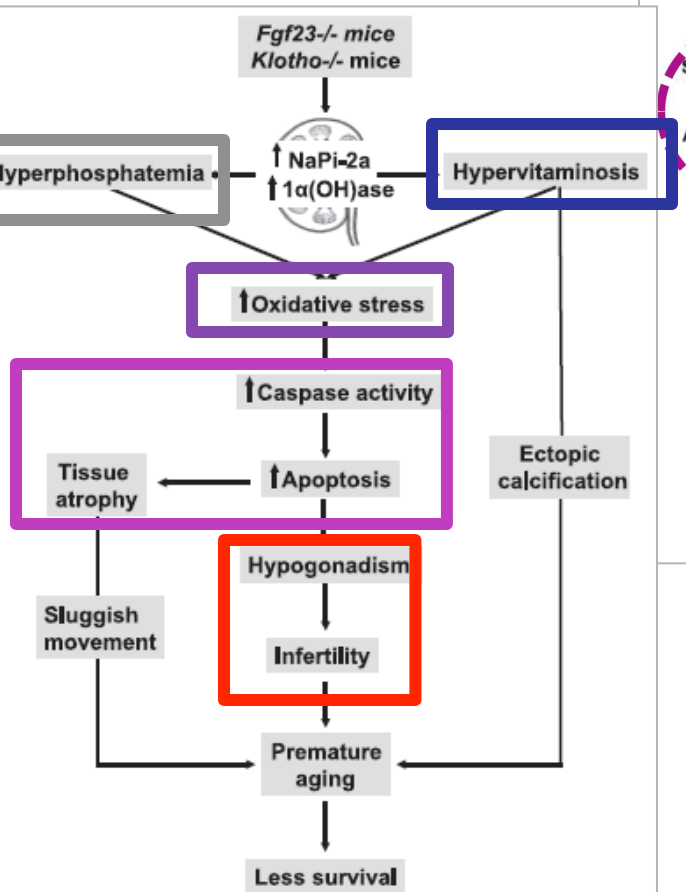
Validation

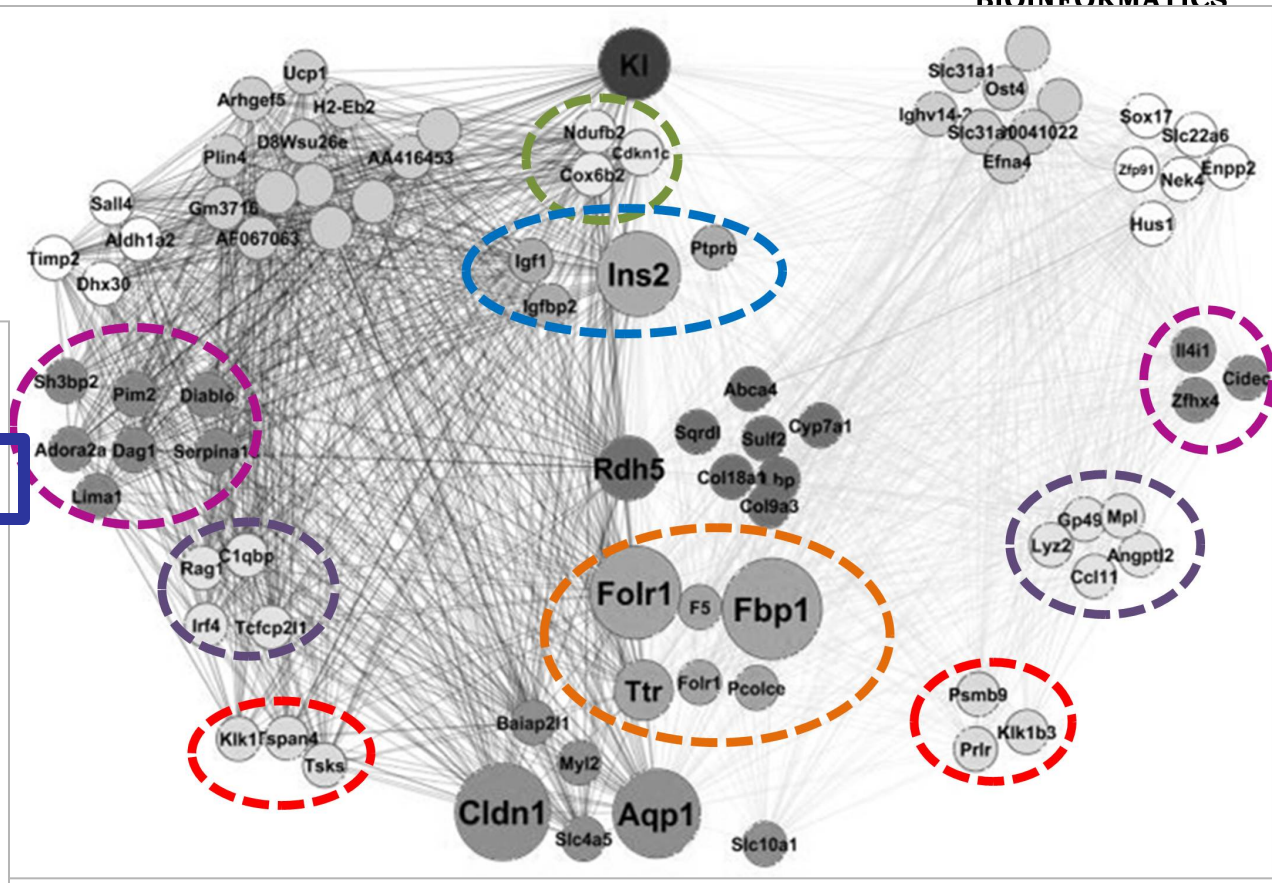




Chateau *et al.* 2010. *Aging*

Subsystem Validation

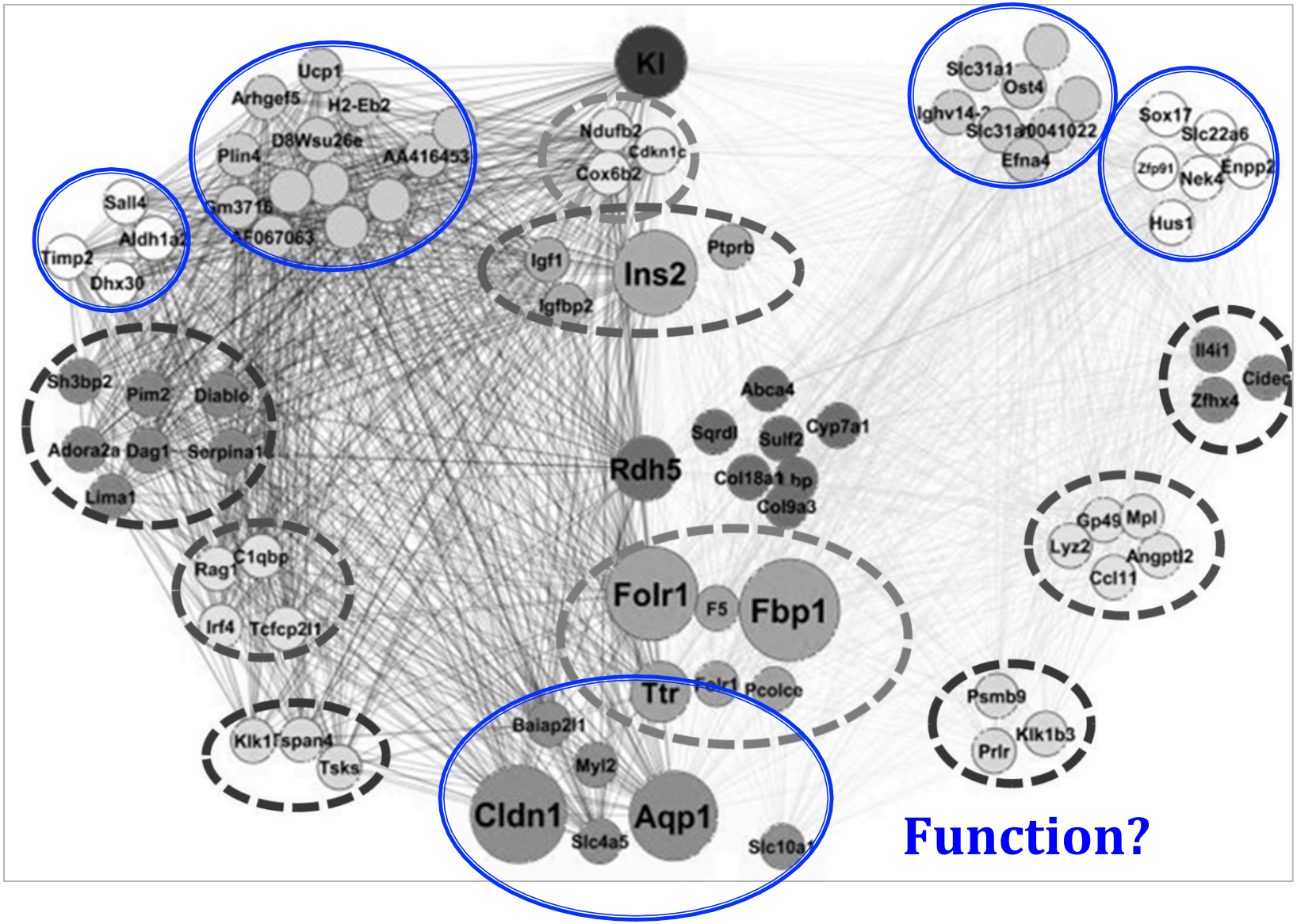




Razzaque , 2009. Am J Physiol Renal Physiol. 296(3):F470-6.

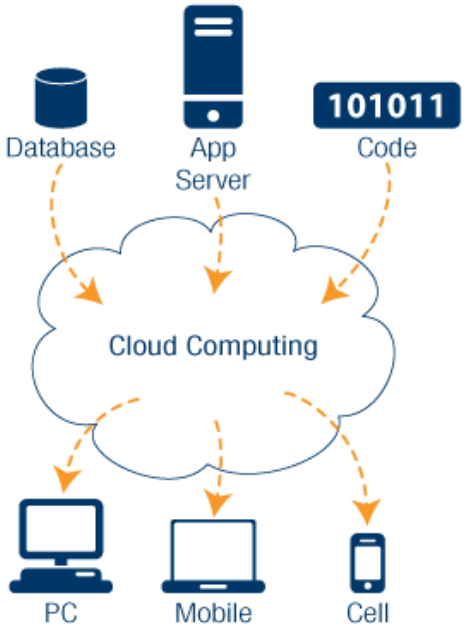
Discovery

UNO BIOINFORMATICS



Working in the Cloud

- Cloud computing is Web-based processing and storage. Software and equipment are offered as a service over the Web.
  - Data and applications can be accessed from any location
  - Data and applications can easily be shared through a common platform
  - Clouds need not be public; companies can introduce private cloud computing solutions

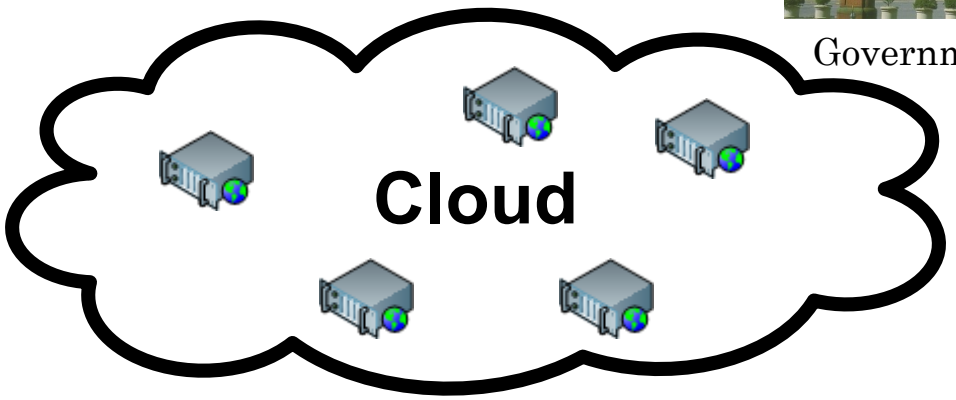


Cost Reduction & Convenience

Small Business

Government Offices

Multinational Corporations



Homes

- Flexible availability of resources
- Opportunity for developers to easily push their applications
- Targeted advertising
- Easy Software Upgrades for customers
  - Example: Webmail

Private Clouds

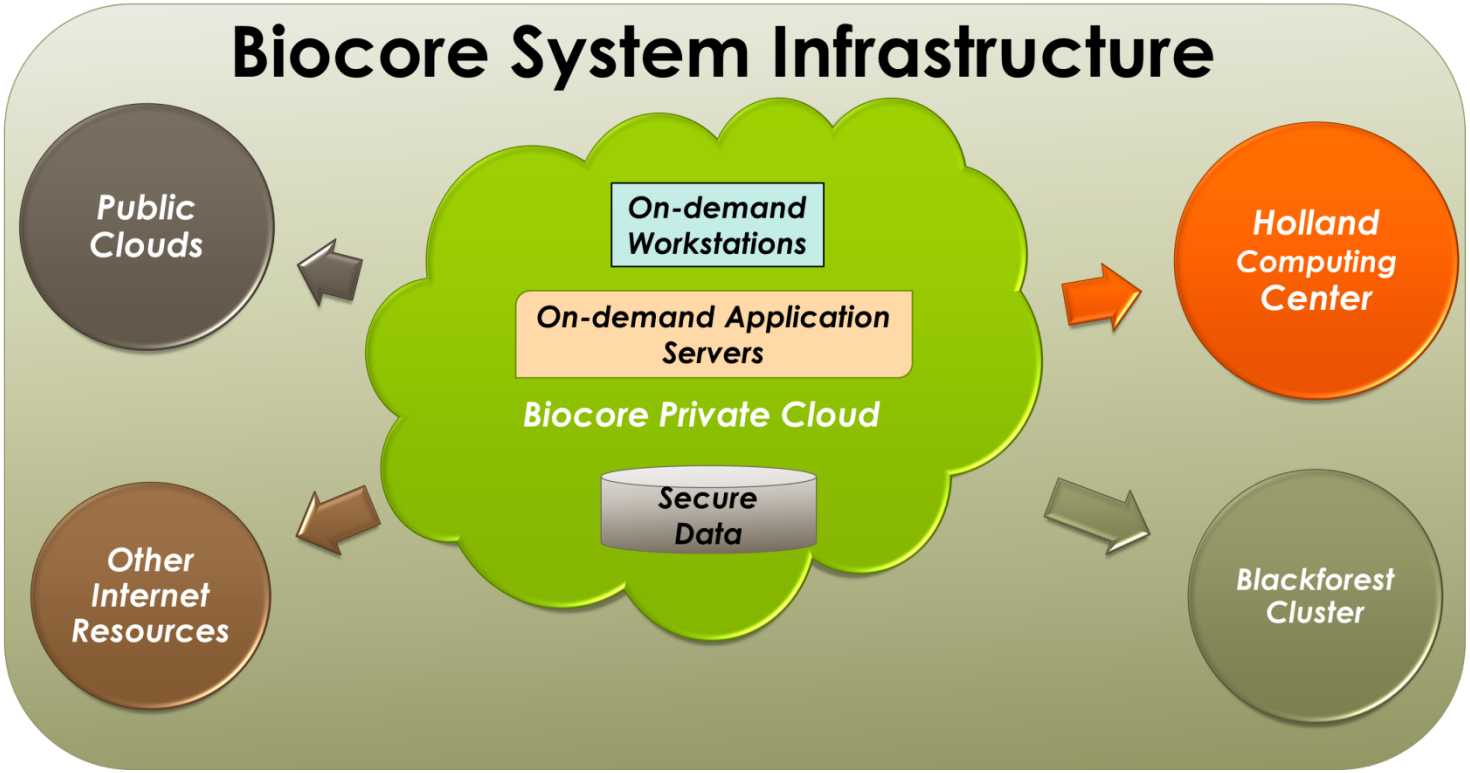
- Core facilities need to acquire private infrastructure-level virtual cloud technology. Best vendor for such technology is VMware. The Bioinformatics Core facility at UNO uses *VMware vSphere Enterprise.*
- Public Clouds like Amazon EC2, RackSpace cannot be used in all cases due to various restrictions put forth by regulations (e.g. HIPAA data locality requirement). Such public clouds could only be used as a scalable platform for already anonymized data.
- Private Virtual Cloud is on-premise solution allowing all the benefits of virtualization technology both from an administrative and end-user perspective.

Security Hardened OS Template

- Virtual layer allows launching servers/machines from a template tested and hardened by the designated Security Office of the organization.

- Easier to isolate security issues on a virtual machine, resolve the issue on a clone of the virtual machine and deploy clone back to production environment.

Proposed Model



Are the current models sufficient to predict the druggable and therapeutic targets that we need?

Eduardo Campos dos Santos
edu@edusantos.eti.br

Issues in new drug development

Challenging paradigm

basic discovery research → funded mainly by → government and philanthropic organizations

Gap: proof of concept

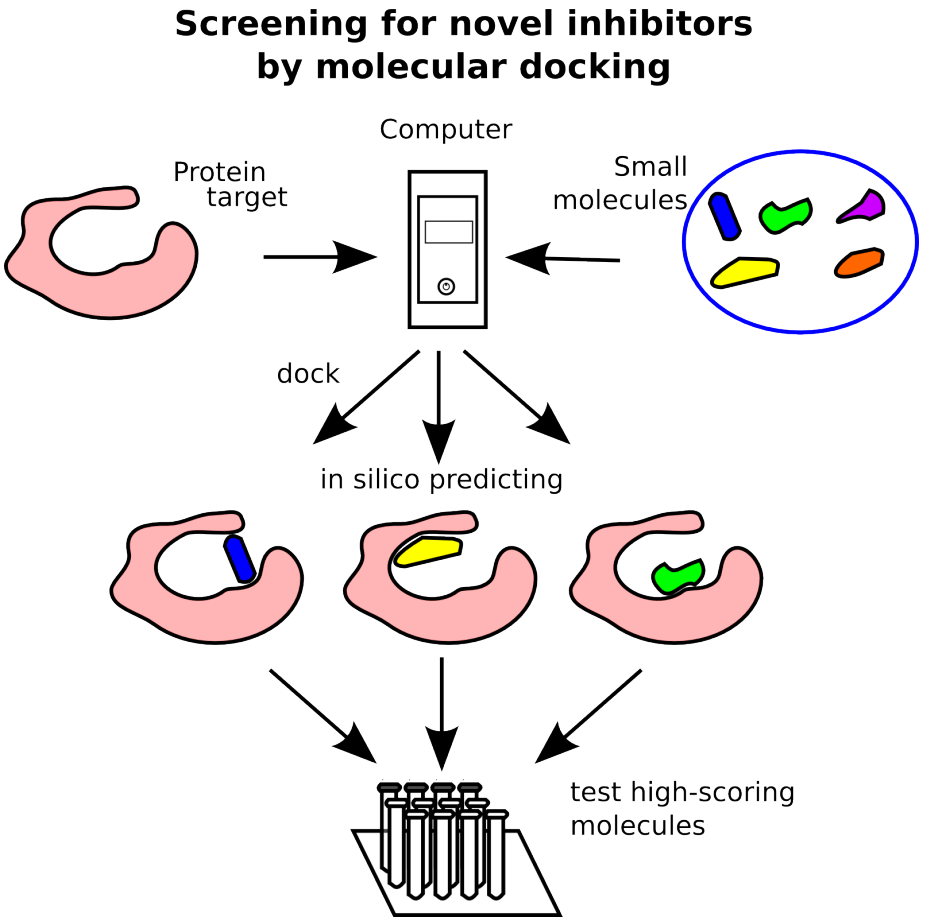
late-stage development → funded mainly by → pharmaceutical companies and venture capitalists

Challenging areas

Cancer treatment →
- hard targets
- to attack cancer cells but to leave healthy cells largely untouched

Neglected diseases →
- low investment
- acquirement of drug resistance

Current methods



Drawback: virtual screening needs to know the protein structure (x-ray crystallography or theoretical by homology)

How to find potential therapeutic targets?
- druggable
- disease related
- specificity

The Druggable Genome
- estimates: not much more (based on current small-molecule drug design)
- do not consider future breakthroughs in medicinal chemistry or biology

Advantage of Using Applied Probability Theory in Bio-tenchonolgy

Hiroshi Toyoizumi

Waseda University

Biotechno2013 (Lisbon)

Advantage of Using Applied Probability Theory

Why math for Bio?

- Bio: real, complex, interesting
- Math: virtual, simple, beautiful
- Business: real, complex, practical

Advantage of Using Applied Probability Theory

Why math for Bio?

- Bio: real, complex, interesting
- Math: virtual, simple, beautiful
- Business: real, complex, practical

Advantage of Using Applied Probability Theory

Why math for Bio?

- Bio: real, complex, interesting
- Math: virtual, simple, beautiful
- Business: real, complex, practical

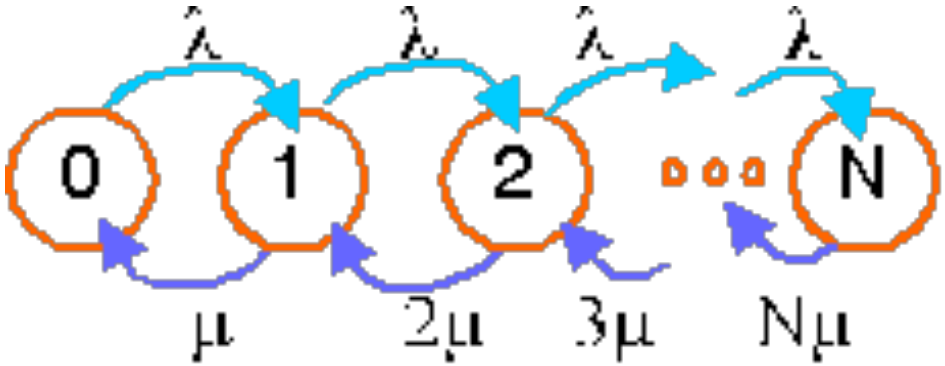
Bio → Math → Business

Erlang-B formula

- Bio: Population
- Math: Birth and Death process (Random Walk)
- Business: Erlang-B formula for designing telephone networks;

$$B = \frac{\rho^N / N!}{\sum_{k=0}^{N} \rho^k / k!}, \tag{1}$$

where $\rho = \lambda / \mu$.



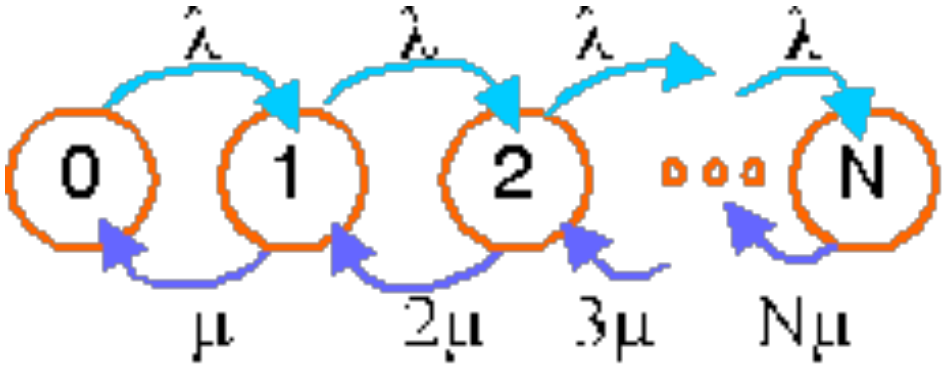
Bio → Math → Business

Erlang-B formula

- Bio: Population
- Math: Birth and Death process (Random Walk)
- Business: Erlang-B formula for designing telephone networks;

$$B = \frac{\rho^N / N!}{\sum_{k=0}^{N} \rho^k / k!}, \tag{1}$$

where $\rho = \lambda / \mu$.



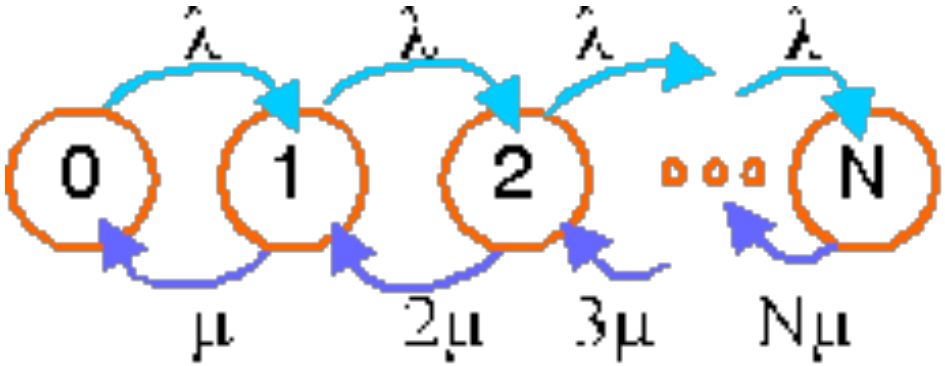
Bio → Math → Business

Erlang-B formula

- Bio: Population
- Math: Birth and Death process (Random Walk)
- Business: Erlang-B formula for designing telephone networks;

$$B = \frac{\rho^N / N!}{\sum_{k=0}^{N} \rho^k / k!}, \tag{1}$$

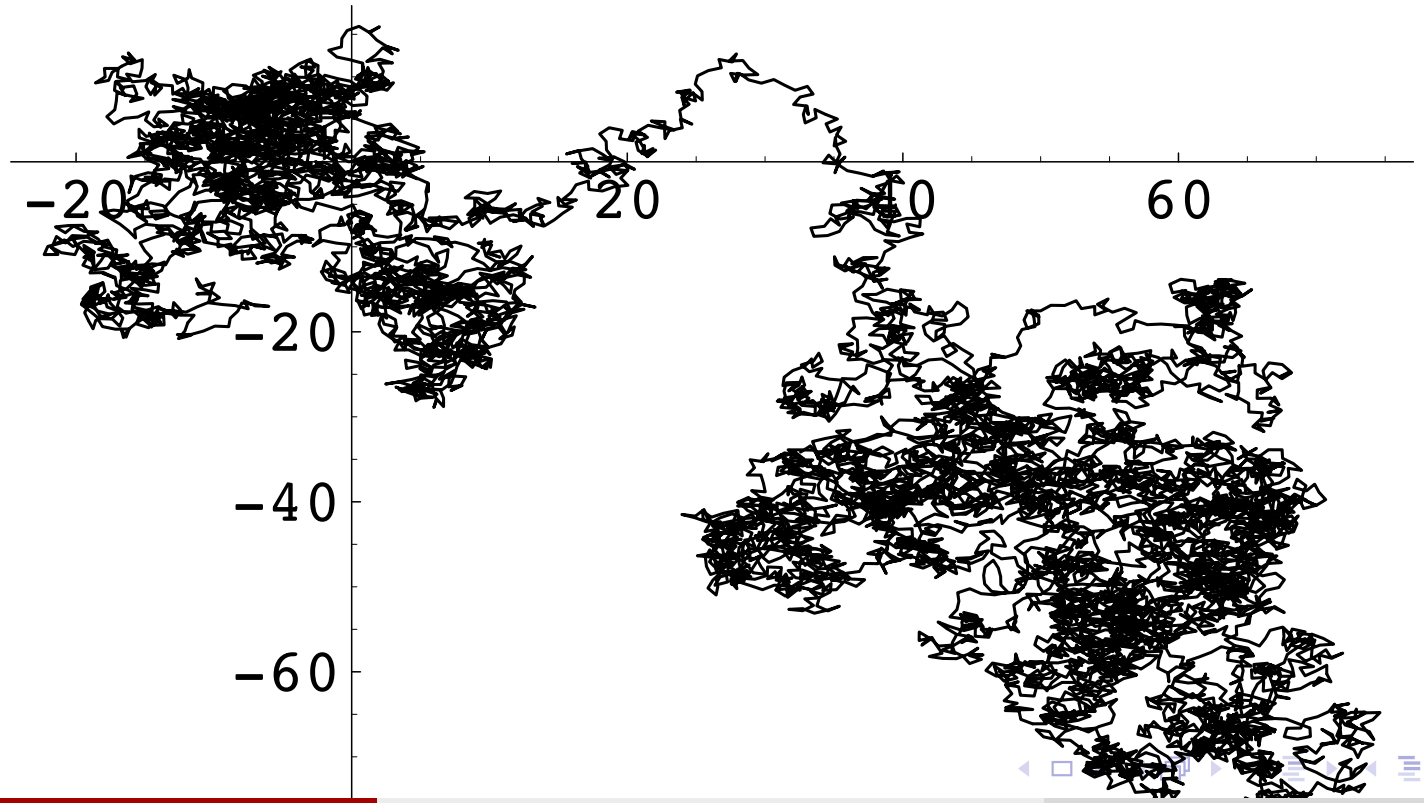
where $\rho = \lambda/\mu$.



Physics → Math → Business

Brownian Motion

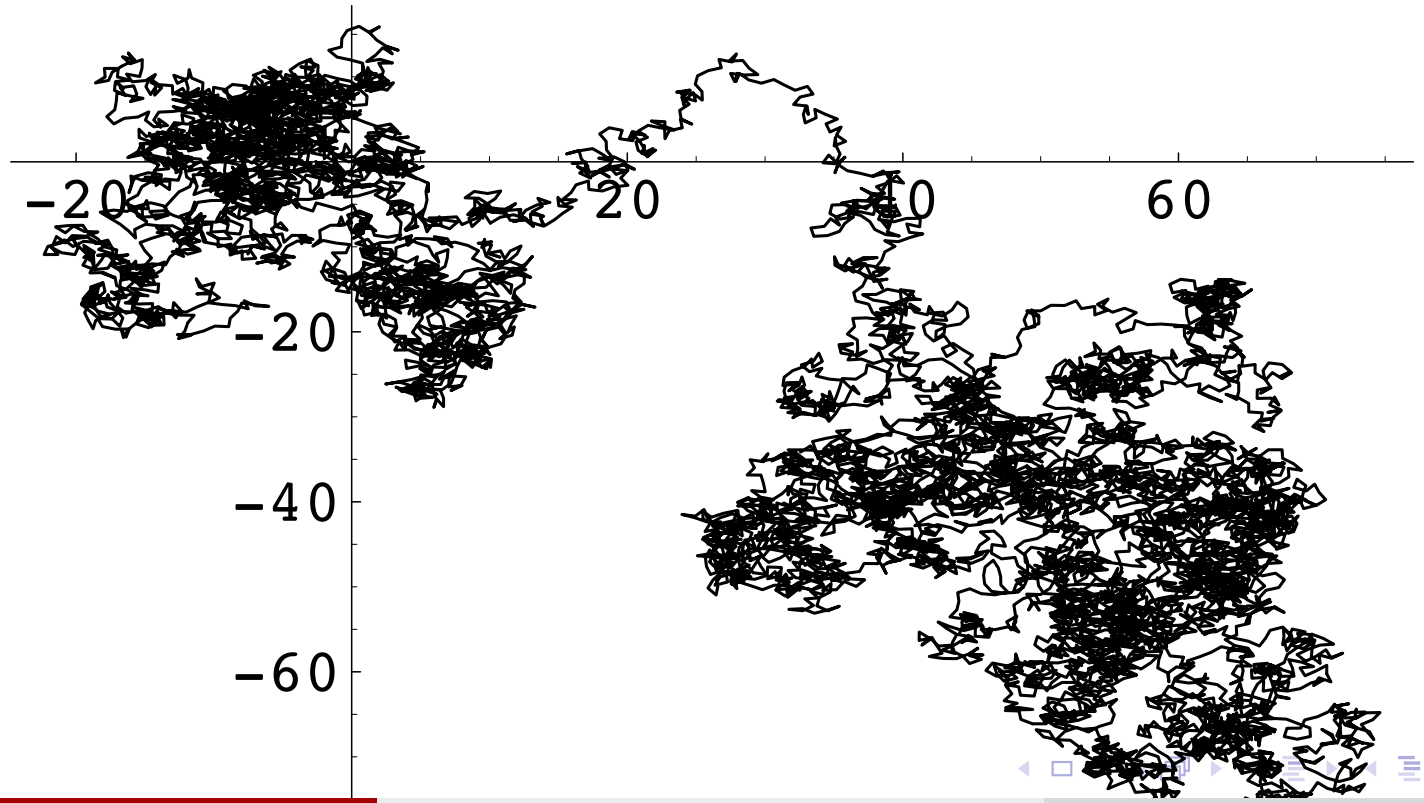
- Physics: Random Movement of Particles
- Math: Brownian Motion $dB(t) = \mu B(t)dt + \sigma dW(t)$.
- Business: Black-Scholes Equation $P = e^{-rt}E[(B(t) - K)^+]$.



Physics → Math → Business

Brownian Motion

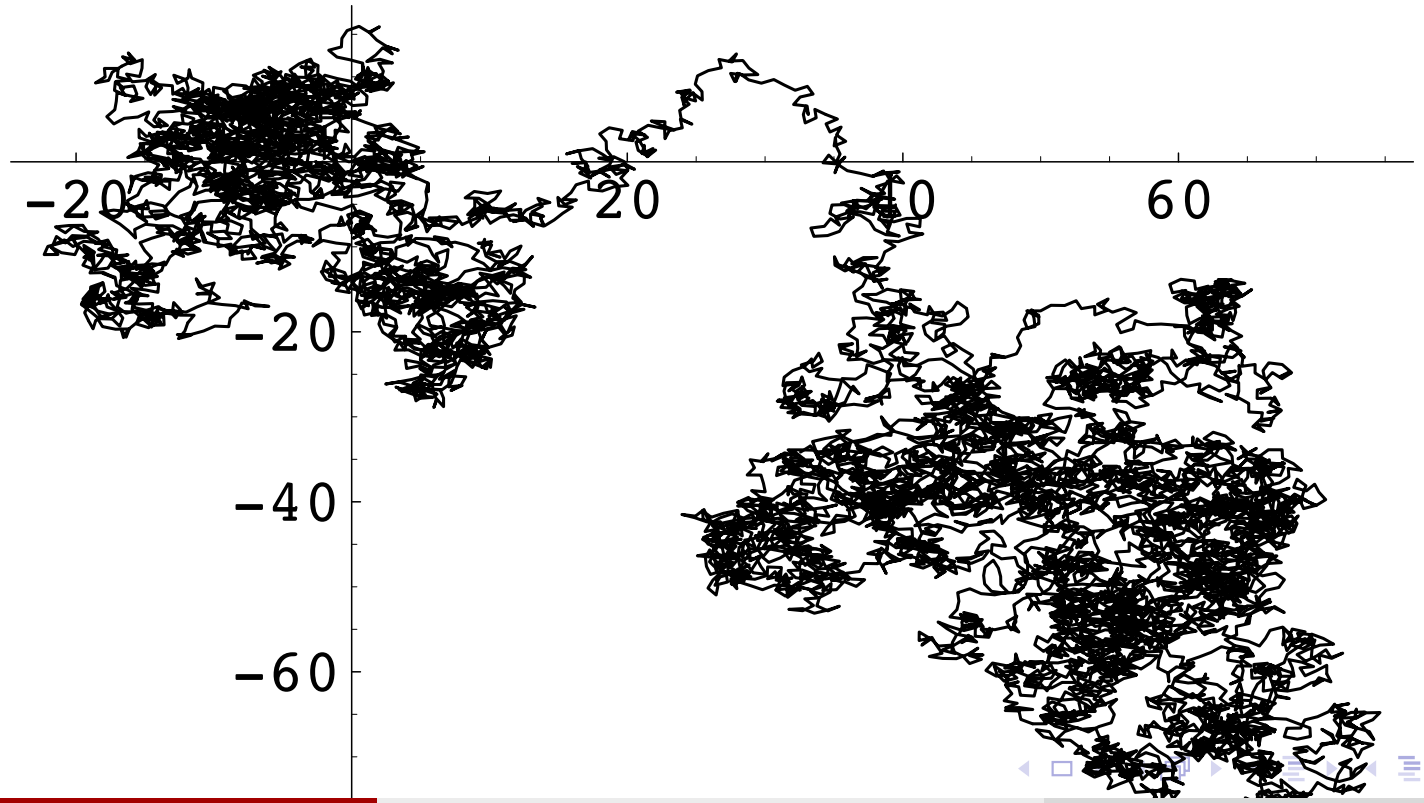
- Physics: Random Movement of Particles
- Math: Brownian Motion $dB(t) = \mu B(t)dt + \sigma dW(t)$.
- Business: Black-Scholes Equation $P = e^{-rt}E[(B(t) - K)^+]$.



Physics → Math → Business

Brownian Motion

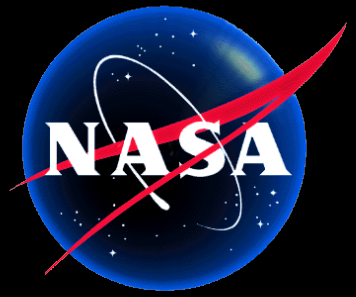
- Physics: Random Movement of Particles
- Math: Brownian Motion $dB(t) = \mu B(t)dt + \sigma dW(t)$.
- Business: Black-Scholes Equation $P = e^{-rt}E[(B(t) - K)^+]$.



Bionature, Lisbon, Portugal, 27 March 2013

PANEL BIO: Are the Current BIO-models Powerful Enough?

Malaria Model and Ringed Seal Research



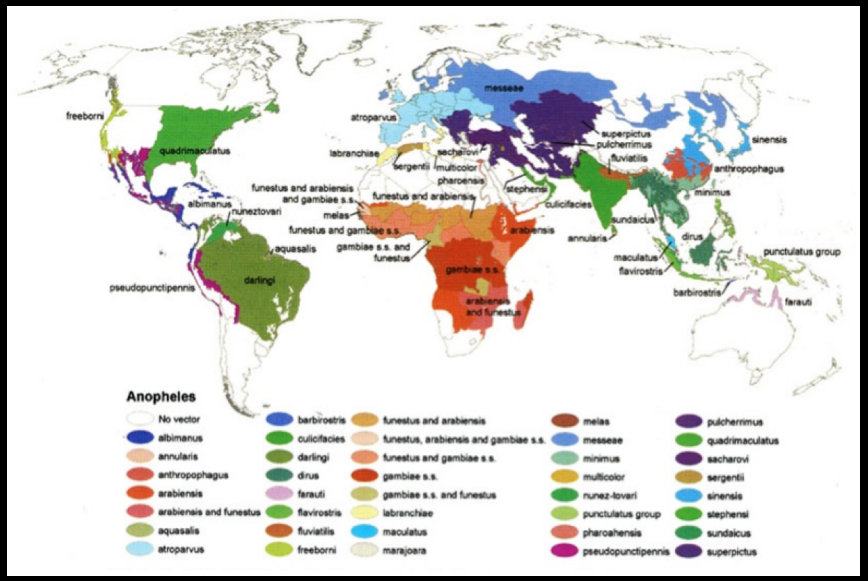
Panelist: Son Nghiem

Jet Propulsion Laboratory

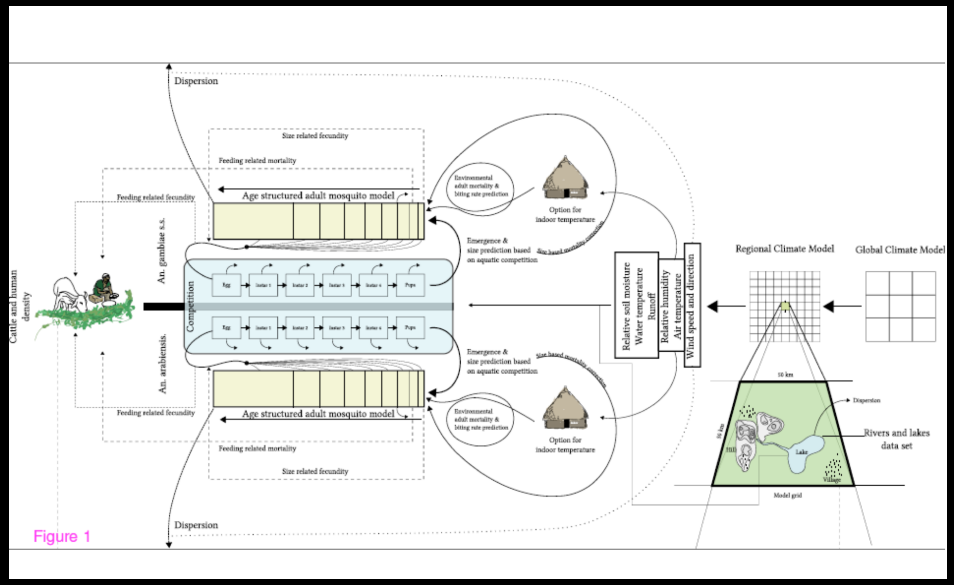
California Institute of Technology,

Pasadena, California, USA

Malaria Model

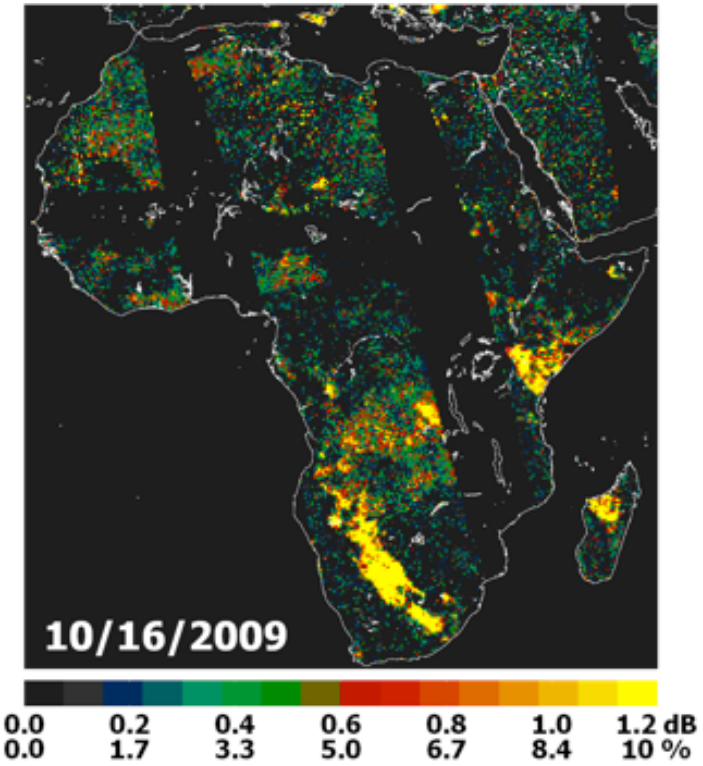


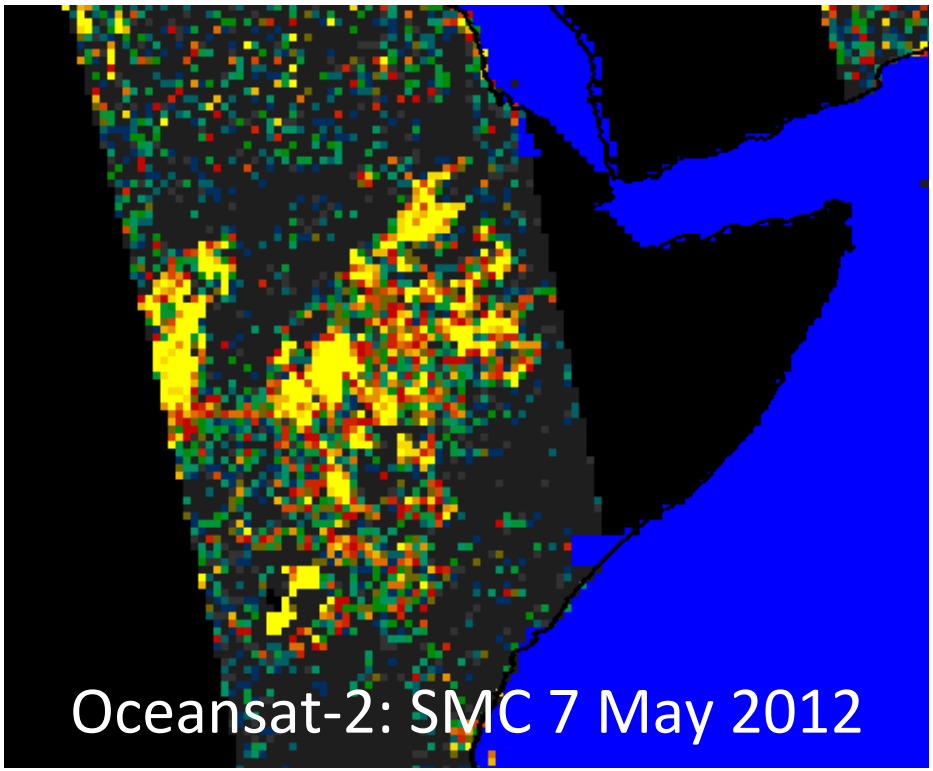
Global distribution of dominant or potentially important malaria vectors (Kiszewksi et al., 2004).



Idealized puddle model (Lunde et al., 2013).

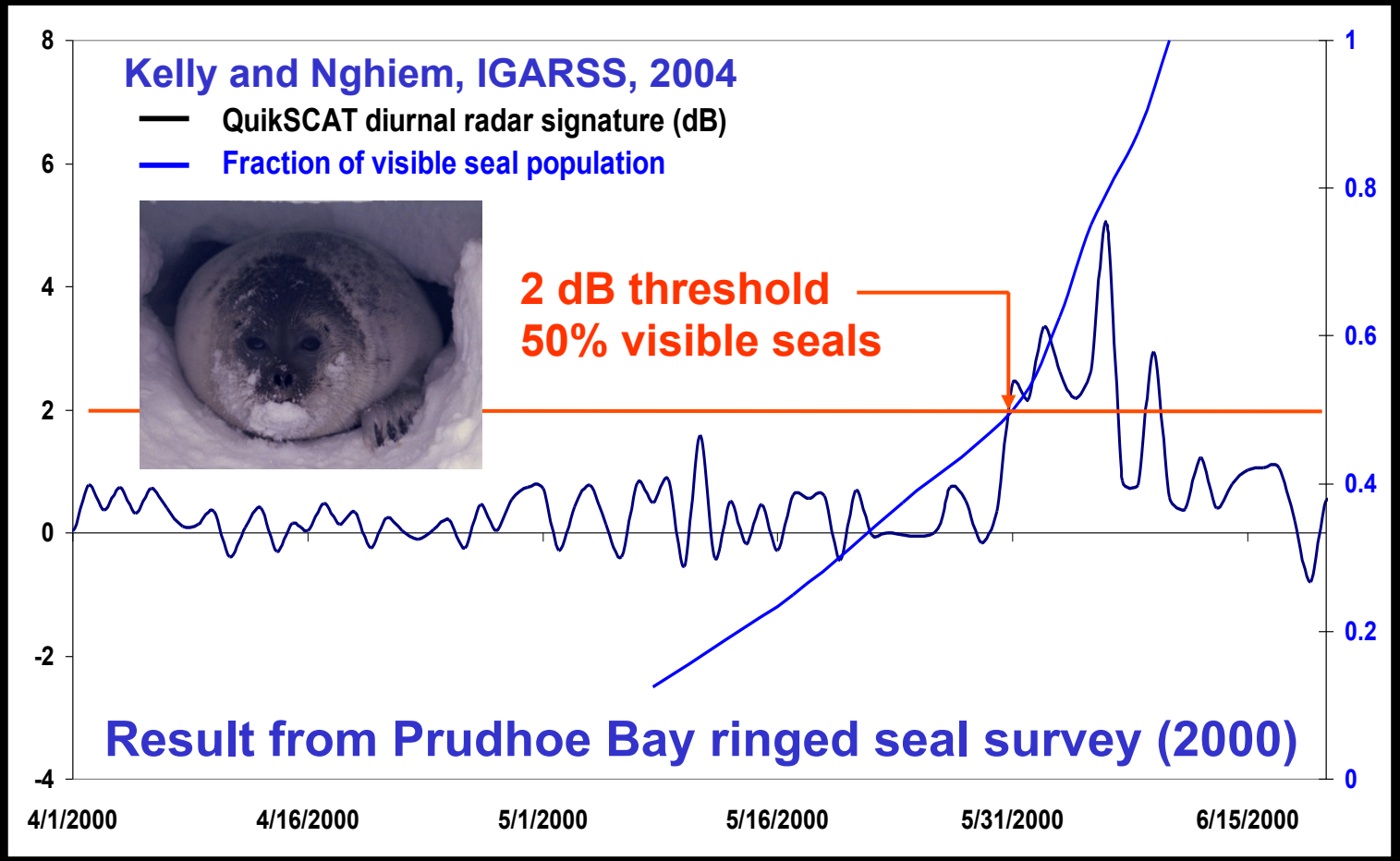
Malaria Model





(a) Left panel: Soil moisture change (*SMC*) across Africa [*Nghiem*, 2009] measured by QuikSCAT with the vertical polarization along ascending orbits on 10/16/ 2009. The color scale represents backscatter change in dB, and volumetric *SMC* in % with the Lonoke rating. Data gaps between orbits are seen in dark bands on land. High *SMC* is observed in an extensive pattern (yellow areas) across South Africa curving northward to Botswana, Namibia, Zambia, and Mozambique, with a wet region seen in Madagascar. (b) Right panel: *SMC* across Ethiopia measured by Oceansat-2 on 5/7/2012. Here we use data only from ascending orbit in this example (all orbits can be used so that the orbit data gaps are small).

Arctic Ringed Seal Research



What if seals come out too early due to early snow melt on sea ice?

Contact

Son.V.Nghiem@jpl.nasa.gov