Highlights of the NIH Common Fund National Centers for Biomedical Computing (NCBC)

July 16, 2012

Presented at ISMB2012 Long Beach, CA

Overview of the NIH Common Fund National Centers for Biomedical Computing (NCBC)

The National Centers for Biomedical Computing (NCBC) are intended to be part of the national infrastructure in Biomedical Informatics and Computational Biology. There are eight Centers (two now retired) that cover biophysical modeling, biomedical ontologies, information integration, tools for gene-phenotype and disease analysis, systems biology, image analysis, and health information modeling and analysis. The centers create innovative software programs and other tools that enable the biomedical community to integrate, analyze, model, simulate, and share data on human health and disease. Each Center has Cores that are focused on (i) biomedical computational science and (ii) driving biological projects (DBPs) whose intent is to foster interaction between computational and biomedical research that leads to innovation in both areas. There are numerous education and training activities sponsored by the Centers, ranging from webinars to tutorials to graduate internships. There is collaboration among the Centers, and also between the Centers and other large NIH program such as the CTSA centers. There is a special grant program for Collaborations with National Centers for Biomedical Computing. Under the collaboration grant program, 225 applications for funding have been reviewed at NIH and 33 have been awarded since its inception 8 years ago. The NCBCs have a federated approach to outreach and dissemination, Biositemaps provides access to unique software repositories and resources developed at the eight Centers. The NCBC Portal http://www.ncbcs.org/ provides a gateway to information about each Center and its products. The Centers work together in order to advertise and promote usage of their tools and resources. For example, a recent special session at the mainline Intelligent Systems in Molecular Biology (ISMB) meeting 2012 in Long Beach on July 16 focused on NCBC accomplishments; and a set of special articles in Journal of the American Medical Informatics Association (JAMIA) has been published.

Three Themes covered by eight grants:

- THEME A: Systems Computational Science
- THEME B: Imaging
- THEME C: Biomedical Informatics

THEME A: Computational biology and systems science under the NCBC program

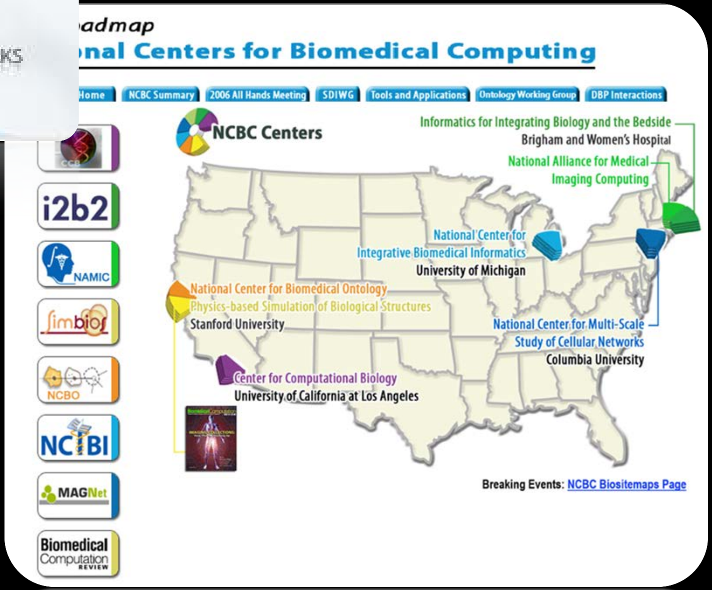
Russ B Altman, MD, PhD
Stanford University
Simbios http://simbios.stanford.edu/

Three NCBC Centers

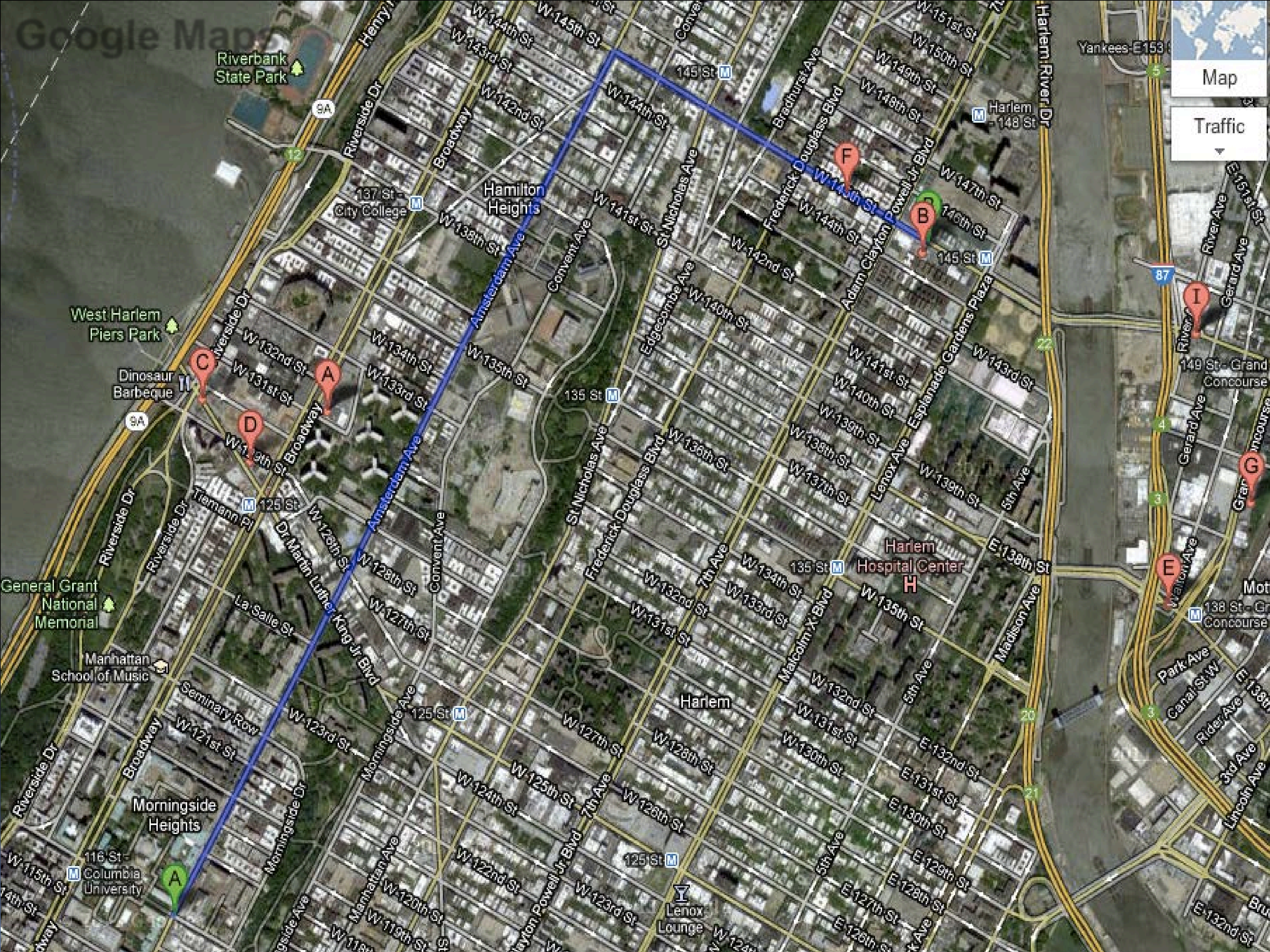
- MAGNet: Multiscale analysis of genomic and cellular networks (Columbia U.. Califano)
- NCIBI: National Center for Integrative Biomedical Informatics (U. Michigan, Athey)
- Simbios: National Center for Physics Based Simulation of Biological Structure (Stanford U., Altman, Delp, Pande)

MAGNet MULTISCALE ANALYSIS OF GENOMIC AND CELLULAR NETWORKS

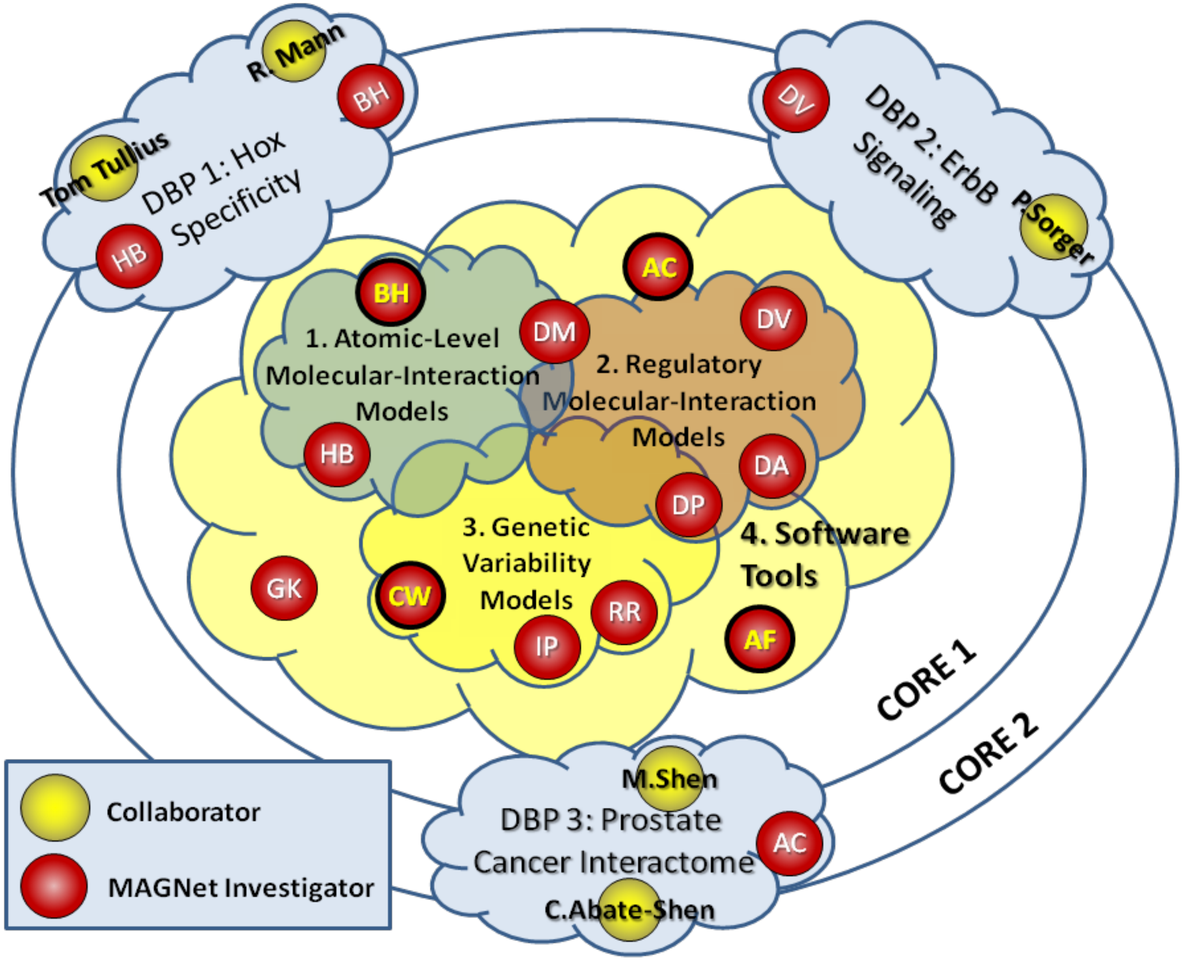
NCBC and ICBP/CCSB Center



A SYSTEMS AND STRUCTURAL BIOLOGY APPROACH TO THE DISSECTION OF COMPLEX BIOLOGICAL PHENOTYPES


Google Maps
Map
Traffic
Riverbank State Park
9A
12
Riverside Dr
Broadway
Henry Hudson Pkwy
W 145th St
W 144th St
W 143rd St
W 142nd St
145 St
Convent Ave
Bradhurst Ave
Frederick Douglass Blvd
Adam Clayton Powell Jr Blvd
W 151st St
W 150th St
W 149th St
W 148th St
W 147th St
W 146th St
W 141st St
Harlem - 148 St
Harlem River Dr
Yankees-E 153 St
5
137 St - City College
Hamilton Heights
W 138th St
Amsterdam Ave
St Nicholas Ave
Edgecombe Ave
W 140th St
Lenox Ave
Esplanade Gardens Plaza
87
River Ave
Gerard Ave
E 151st St
149 St - Grand Concourse
W 143rd St
22
West Harlem Piers Park
W 132nd St
W 131st St
W 134th St
W 133rd St
W 135th St
135 St
W 136th St
W 139th St
W 137th St
5th Ave
4
Dinosaur Barbeque
W 129th St
125 St
Tiemann Pl
Dr Martin Luther King Jr Blvd
W 128th St
W 127th St
Convent Ave
Grand Concourse
3
Harlem Hospital Center
E 138th St
Walton Ave
138 St - Grand Concourse
Mott Haven
General Grant National Memorial
La Salle St
7th Ave
Malcolm X Blvd
Madison Ave
Park Ave
Canal St W
Manhattan School of Music
Seminary Row
W 123rd St
W 121st St
Harlem
20
Rider Ave
3rd Ave
E 132nd St
E 131st St
E 130th St
E 129th St
E 128th St
E 127th St
E 126th St
W 125th St
W 124th St
W 122nd St
W 120th St
W 119th St
Morningside Ave
Manhattan Ave
Morningside Heights
Morningside Dr
116 St - Columbia University
W 115th St
W 114th St
Lenox Lounge
Lincoln Ave
21
A
B
C
D
E
F
G
I

The strategy of MAGNet



Leadership:
Andrea Califano (PI)
Barry Honig (co-PI)
Aris Floratos (Exec Dir)

Investigators:
Dimitris Anastassiou
Harmen Bussemaker
Gail Kaiser
Diana Murray
Dana Pe’er
Itsik Pe’er
Raul Rabadan
Dennis Vitkup
Chris Wiggins

Driving Biological Projects
Cory Abate-Shen
Michael Shen
Richard Mann
Peter Sorger
Tom Tullius

MAGNet Center Mission: the development of novel Structural and Systems Biology methods and tools for the *dissection of molecular-interactions in the cell* and for the *interaction-based elucidation of cellular phenotypes*

Core I: Theoretical and Computational Studies

- ***Project 1: Atomic-level Molecular-Interaction Models***
  - B.Honig, H.Bussemaker, D.Murray
  - structural and biophysical properties of proteins and DNA and for their integration into Systems Biology.
- ***Project 2: Regulatory Molecular-Interaction Models***
  - A.Califano, D.Anastassiou, D.Pe’er, D.Vitkup.
  - The multiscale complexity of intra- and inter-cellular regulatory studied by integrating layers of regulation traditionally studied in isolation.

Core I: Theoretical and Computational Studies

- ***Project 3: Genetic Variability Models***
  - ***C.Wiggins****, D.Pe'er, I.Pe'er, R.Rabadan*
  - Tools for the genotypic data analysis,

- ***Project 4: Software tools***
  - **A.Floratos**, A.Califano, G.Kaiser
  - Algorithms developed in Projects 1 – 3 will be integrated and disseminated as *geWorkbench* platform components.

Driving Biological Projects

- ***DBP 1: Role of DNA shape in protein DNA binding specificity***
  - **T.Tullius**, R.Mann, H.Bussemaker, B.Honig
  - Determine the sequence and structural basis of Hox protein specificity.

- ***DBP 2: Probabilistic dynamical models of ErbB signaling pathways***
  - ***P. Sorger*** *and D. Vitkup.*
  - Predictions of network-mediated effects of oncogenic mutations and small-molecule perturbations
- **DBP 3: Master regulators of tumorigenesis and drug-sensitivity in prostate cancers,**
  - ***C.Abate-Shen****, M.Shen, A.Califano*
  - Human prostate cancer xenografts in the mouse to reconstruct *in vivo* regulatory networks and elucidate the mechanisms driving tumorigenesis and drug sensitivity.

Accomplishments: Science, Infrastructure, Education

Publications:

- 61 MAGNet publications (last 2Y)
  - >50% in high impact factor journals (18/28 in the last year)

Collaborative Projects:

- 10 Driving Biological Projects
- >60 collaborative projects
  >1/3 funded

Infrastructure

- 4,500 CPU supercomputer, ~1 petabyte storage system
- Scored 23 on a $2M equipment grant to upgrade our supercomputer

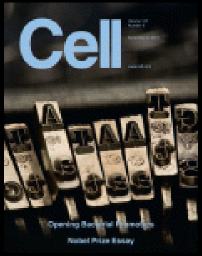
Education and Dissemination

- Conference organization:
  - 2012 AACR Special Conference on Chemical Systems Biology
  - 2012 RECOMB Regulatory and Systems Biology Conference with DREAM Challenges
  - NYAS Sys. Bio. Interest Group

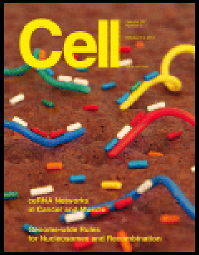
Students and Postdoctoral training

- Over 100 students and postdocs in MAGNet funded labs
- Developed over 20 CompBio courses
- Training Grant (NIGMS)
- Stats from student retention: 6/8 of our top choices decided to come to Columbia

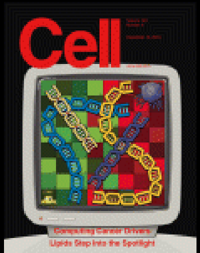
Scientific Productivity



Cofactor Binding Evokes Latent Differences in DNA Binding Specificity between Hox Proteins. Slattery M, Riley T, Liu P, Abe N, Gomez-Alcala P, Dror I, Zhou T, Rohs R, Honig B, Bussemaker HJ, Mann RS. Cell. 2011 Dec 9;147(6)



An Extensive MicroRNA-Mediated Network of RNA-RNA Interactions Regulates Established Oncogenic Pathways in Glioblastoma Sumazin P, Yang X, Chiu HS, Chung WJ, Iyer A, Llobet-Navas D, Rajbhandari P, Bansal M, Guarnieri P, Silva J, Califano A Cell. 2011 Oct 14;147(2):370-81.



An integrated approach to uncover drivers of cancer. Akavia UD, Litvin O, Kim J, Sanchez-Garcia F, Kotliar D, Causton HC, Pochanard P, Mozes E, Garraway LA, Pe'er D. Cell. 2010 Dec 10;143(6):1005-17. Epub 2010 Dec 2.

The transcriptional network for mesenchymal transformation of brain tumours. Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X, Snyder EY, Sulman EP, Anne SL, Doetsch F, Colman H, Lasorella A, Aldape K, Califano A, Iavarone A. Nature. 2010 Jan 21;463(7279):318-25. Epub 2009 Dec 23.

The role of DNA shape in protein-DNA recognition. Rohs R, West SM, Sosinsky A, Liu P, Mann RS, Honig B. Nature. 2009 Oct 29;461(7268):1248 53.

Geographic dependence, surveillance, and origins of the 2009 influenza A (H1N1) virus. Trifonov V, Khiabanian H, Rabadan R. N Engl J Med. 2009 Jul 9;361(2):115-9.

High Impact Publications (2011 - 2012)

- Chen, Y. … ll, **B. Honig**, H. Lu and W. Cho (2012). "Genome-wide Functional Annotation of Dual-Specificity Protein- and Lipid-Binding Modules that Regulate Protein Interactions." *Molecular Cell*.
- Dapito, …, R. Rabadan and R. F. Schwabe (2012). "Promotion of Hepatocellular Carcinoma by the Intestinal Microbiota and TLR4." *Cancer Cell* **21**(4): 504-516.
- Fabbri, … R. Rabadan, R. Dalla-Favera and G. Gaidano (2011). "Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation." *The Journal of Experimental Medicine* **208**(7): 1389-1401.
- Fischer, M., …, B. Honig and D. Petrey (2011). "MarkUs: a server to navigate sequence-structure-function space." *Nucleic Acids Res* **39**(Web Server issue): W357-361.
- Greenbaum, … R. Rabadan (2012). "Viral reassortment as an information exchange between viral segments." *PNAS* **109**(9): 3341-3346.
- Grossmann, … R. Rabadan, T. Haferlach and B. Falini (2011). "Whole-exome sequencing identifies somatic mutations of BCOR in acute myeloid leukemia with normal karyotype." *Blood* **118**(23): 6153-6163.
- Kreimer, A.… D. Pe'er and I. Pe'er (2012). "Inference of modules associated to eQTLs." *Nucleic Acids Res*.
- Martinez, M. R., … A. Califano and G. A. Stolovitzky (2012). "Quantitative modeling of the terminal differentiation of B cells and mechanisms of lymphomagenesis." *PNAS* **109**(7): 2672-2677.
- Ntziachristos, … R. Rabadan, et al. (2012). "Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia." *Nature Medicine* **18**(2): 298-301.

High Impact Publications (2011 - 2012)

- Pasqualucci, L., … R. Rabadan and R. Dalla-Favera (2011). "Analysis of the coding genome of diffuse large B-cell lymphoma." *Nature Genetics* **43**(9): 830-837.
- Rossi, D., … R. Rabadan, et al. (2011). "Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness." *Blood* **118**(26): 6904-6908.
- Rossi, D., … R. Rabadan, et al. (2012). "Mutations of NOTCH1 are an independent predictor of survival in chronic lymphocytic leukemia." *Blood* **119**(2): 521-529.
- Slattery, B. Honig, H. J. Bussemaker and R. S. Mann (2011). "Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins." *Cell* **147**(6): 1270-1282.
- Sumazin, … and A. Califano (2011). "An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma." *Cell* **147**(2): 370-381.
- Tiacci, … R. Rabadan and B. Falini (2011). "BRAF mutations in hairy-cell leukemia." *The New England Journal of Medicine* **364**(24): 2305-2315..
- Zhang, …, B. Honig and D. Petrey (2011). "PredUs: a web server for predicting protein interfaces using structural neighbors." *Nucleic Acids Res* **39**(Web Server issue): W283-287.
- Della Gatta, … A. Califano and A. A. Ferrando (2012). "Reverse engineering of TLX oncogenic transcriptional networks identifies RUNX1 as tumor suppressor in T-ALL." *Nat Med* **18**(3): 436-440.
- Pasqualucci, … R. Rabadan and R. Dalla-Favera (2011). "Analysis of the coding genome of diffuse large B-cell lymphoma." *Nature Genetics* **43**(9): 830-837.

MAGNet derived activities and links

New Department of Systems Biology at Columbia University

- 5 New Recruitments in the last year (27 faculty)
  - Saeed Tavazoie, Ph.D (Princeton)
  - Sagi Shapira, Ph.D (Broad)
  - Chaolin Zhang, Ph.D (Rockefeller)

  Peter Simms, Ph.D (Harvard)

  Yufeng Shen, Ph.D (Columbia)
- Integration of HT Biology (NextGen Sequencing + HT Screening) and Computational Biology

8 Centers of Excellence + Several multiPI grants

- National Center for Biomedical Computing (also ICSB/CCSB)
- PSI Initiative (Bioinformatics Core)
- Cancer Target Discovery and Development (CTD^2)
- LINCS Technology Center
- LINCS Computational Center
- In Silico Research Center of Excellence (temporarily terminated)
- Serious Adverse Events Consortium, Data Analysis and Coordination Center
- caBIG MAT Knowledge Center (discontinued)

Software Dissemination

- Individual Tools
  - Highly accessed tools for:

    Reverse Engineering of Regulatory Networks: ARACNe, MINDy, MatrixReduce, Medusa, PREPPI, Target Explorer, etc.

    Protein Structure Analysis: Pudge, Delphi, Grasp, PrISM, etc.

    Identification of causal drivers: MARINa, Connexic, Genatomy,

    …

- Integrative Analysis Platform
  - GeWorkbench

    >5,000 downloads

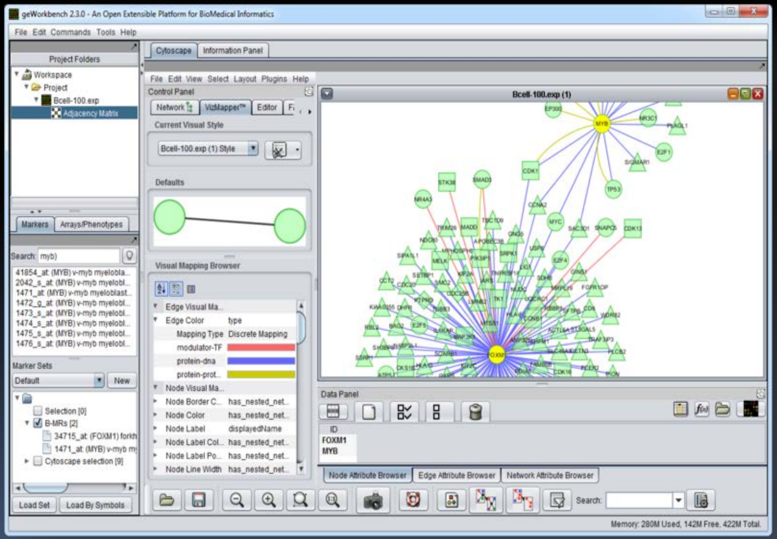
    > 500 regular users

    Built-in Social Science Network

    Large number of functional modules

    Access to most of the software produced by MAGNet investigators

    New version (2012): Web-based user interface



NCIBI: Recent Accomplishments, Growing Community Engagement and Evolving Directions

Brian D. Athey, PI
Jim Cavalcoli, Project Manager
Kevin Smith, tranSMART Acting Exec. Dir.

portal.ncibi.org/gateway & tranSMARTproject.org

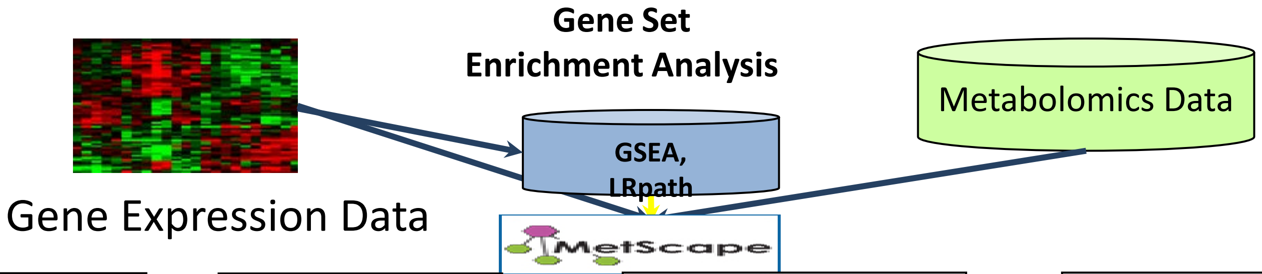
University of Michigan
Health System

NCIBI
National Center for Integrative Biomedical Informatics

The National Center for Integrative Biomedical Informatics (NCIBI) mission is to enhance and accelerate basic, clinical and translational research by providing services to integrate and analyze molecular and phenotypic information

Recent Refinement of NCIBI Metabolomics DBP tools & publications

- Continued effort with MiMI interactions information resources; NICIBi Cytoscape plugins; MetScape 2 and Metab2Mesh Metabolomics tools
  - Key information and analytics platform elements of “Michigan Comprehensive Regional Metabolomics Resource Core”: NIDDK NIH Metabolomics U24 Resource (expected to be launched Sept., 2012)

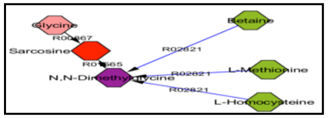


Gene Set Enrichment Analysis

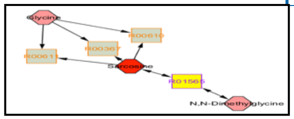
Metabolomics Data

GSEA, LRpath

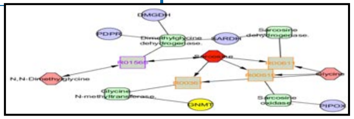
Gene Expression Data



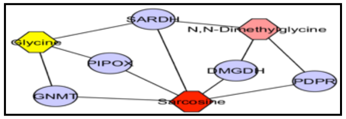
Compound Network



Compound - Reaction Network



Compound - Reaction- Enzyme - Gene Network



Compound - Gene Network

- Karnovsky A et al, **Metscape 2 Bioinformatics Tool for the Analysis and Visualization of Metabolomics and Gene Expression Data** Bioinformatics 2012; 28(3): 373-80. PMID 22135418

  Sartor MA,, et al, **Metab2MeSH: Annotating Compounds with Medical Subject Headings (MeSH)** Bioinformatics 2012; 28(10): 1408-10. PMID 22492643

University of Michigan
Health System

NCIBI
National Center for Integrative Biomedical Informatics

NCIBI services and outreach and training continue

Broad use of web services for data and analytic services (ncibi.ws

- Data Services
  - Natural Language Processing Pipeline for PubMed and PMCOA
  - Gene2MeSH
  - Metab2MeSH
  - Michigan Molecular Interactions Database (MiMI) Protein-Protein Interactions
  - Metabolomics reference data resources from NIH Regional Metabolomics Core (Launching Oct., 2012)
- Computational Analysis Services
  - Natural Language Processing
    - Sentence Segmentation
    - Phrase Structure Parsing
  - Gene Set Enrichment Analysis
    - LRPath
    - ThinkBACK

Outreach activites are growing in frequency and potential impact

- MetScape tutorial (video)
  - http://treehouse.ccmb.med.umich.edu/mst.html
- International Society of Nephrology Forefronts Symposium on Systems Biology and the Kidney
  - Hands-on tutorial session covering NCIBI tools (June, 2012)
- AACR Frontiers in Basic Cancer Research, San Francisco, California - September 14-18, 2011
  - "Lrpath analysis reveals common pathways dysregulated via DNA methylation across cancer types" - Maureen Sartor: NICIB faculty leader
- Poster presentations at GLBIO, May 15, 2012
  - NodeFilter: A Cytoscape Plugin for Efficient Network Exploration

13th RCMI International Symposium on Health Disparities scheduled for December 2012 in San Juan, Puerto Rico(Workshop on NCIBI Tools)

University of Michigan
Health System

NCIBI
National Center for Integrative Biomedical Informatics

NCIBI is transitioning as the organizing leader of the tranSMART Public Private Partnership

- tranSMART is an emerging global open-source public private partnership community developing a comprehensive informatics-based analysis and data-sharing cloud platform for clinical/translational research
- TranSMART Includes pharmaceutical and other companies, not-for-profits, academics, patient advocacy groups, and government stakeholders
- tranSMART Community members collaboratively build, share, and use the tranSMART platform to break down social/technical barriers and capture scientific/collaborative opportunities using crowd sourcing
- NCBCs and tranSMART: tranSMART has i2b2 and functionalities embedded into its code base now. NCIBI to follow October 1, 2012. i2b2 SMART January 1, 2012 release. Preliminary discussions with NCBO.
- NCIBI Core team intact and growing. See tranSMART Project.org

NCIBI
National Center for Integrative Biomedical Informatics

i2b2

SMART

tranSMART
Consortium

Simbios: National Center for Physics-based <u>Sim</u>ulation <u>o</u>f <u>Bio</u>logical <u>S</u>tructure

Russ Altman
Scott Delp
Vijay Pande
Stanford University

Two key scientific foci

- Biomechanical simulation & model sharing (OpenSIM toolkit)

- Molecular dynamics simulation (OpenMM)

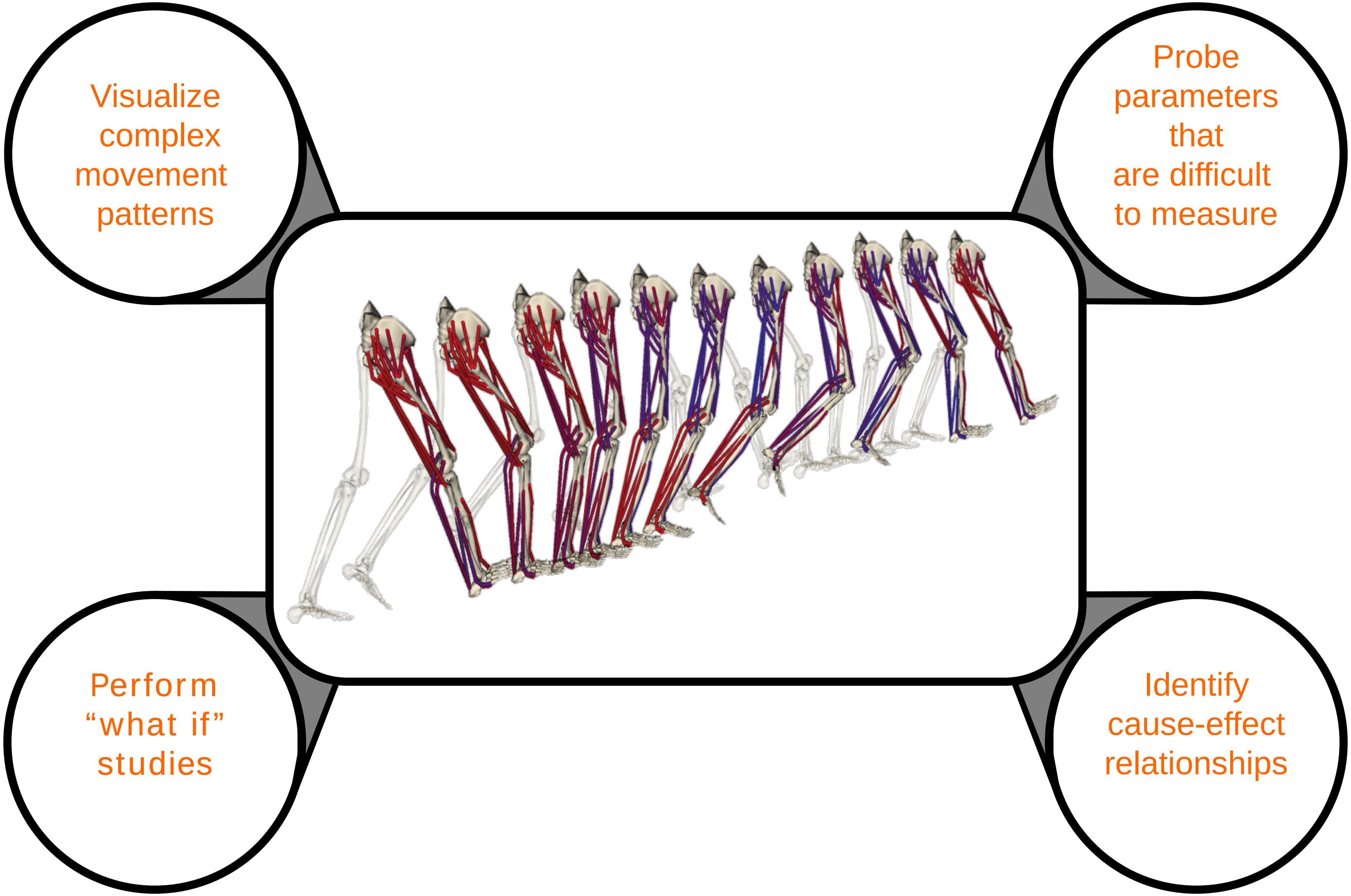
OpenSim: a biologically accurate simulator to evaluate and design Warrior Web technologies

Scott Delp

Departments of Bioengineering,

Mechanical Engineering, Orthopaedic Surgery,
Stanford University

Purpose of modeling and simulation



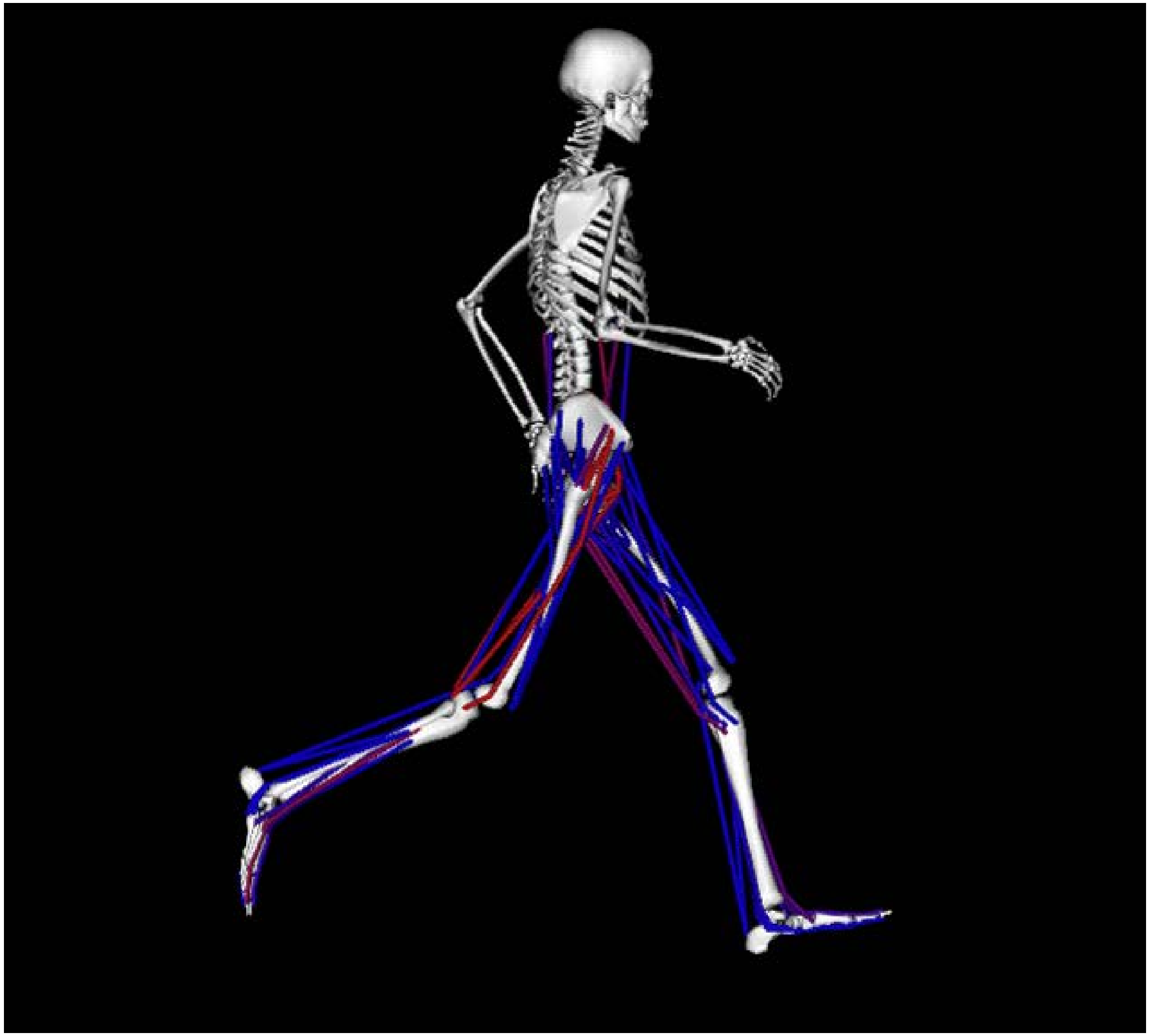
Visualize complex movement patterns

Probe parameters that are difficult to measure

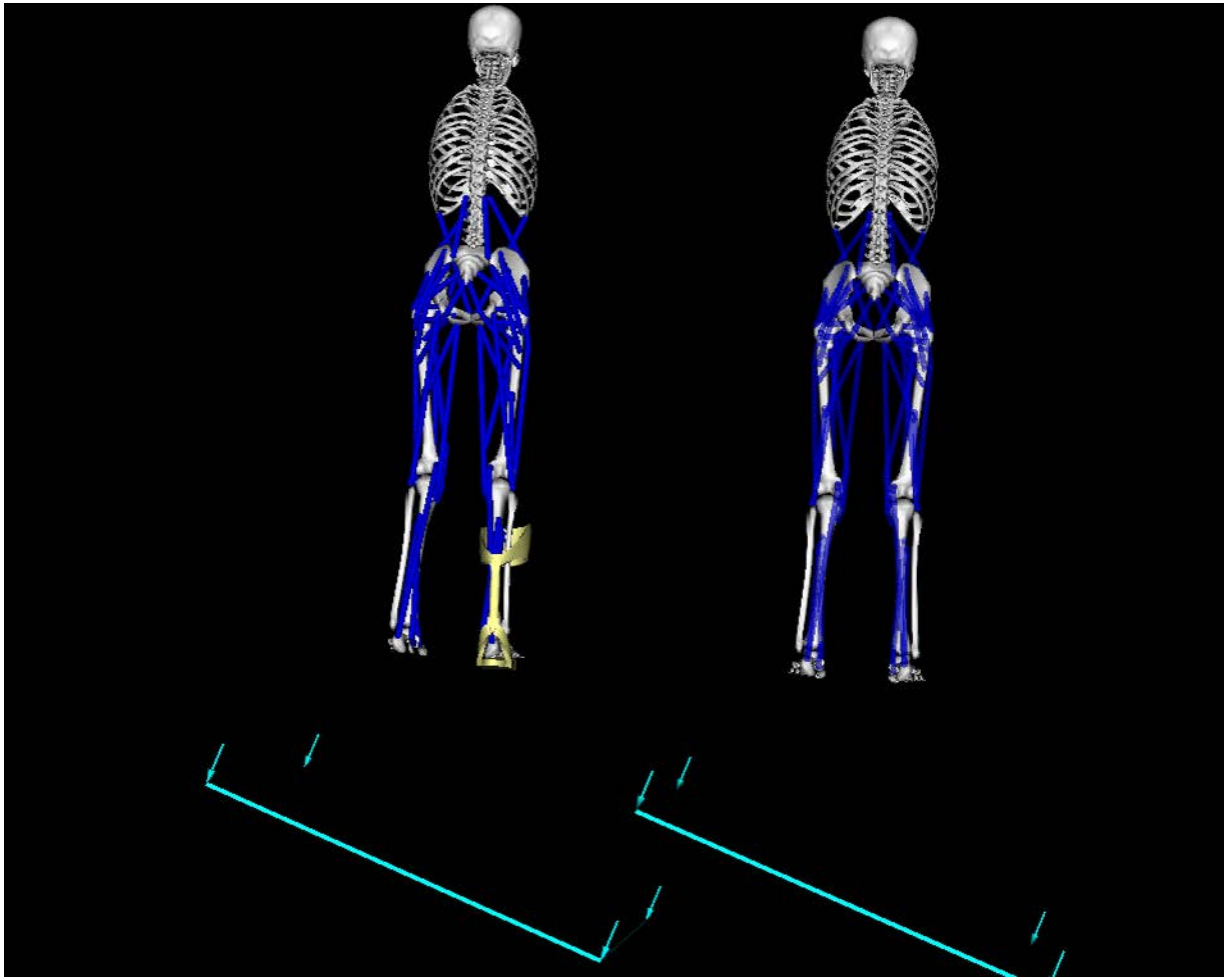
Perform "what if" studies

Identify cause-effect relationships

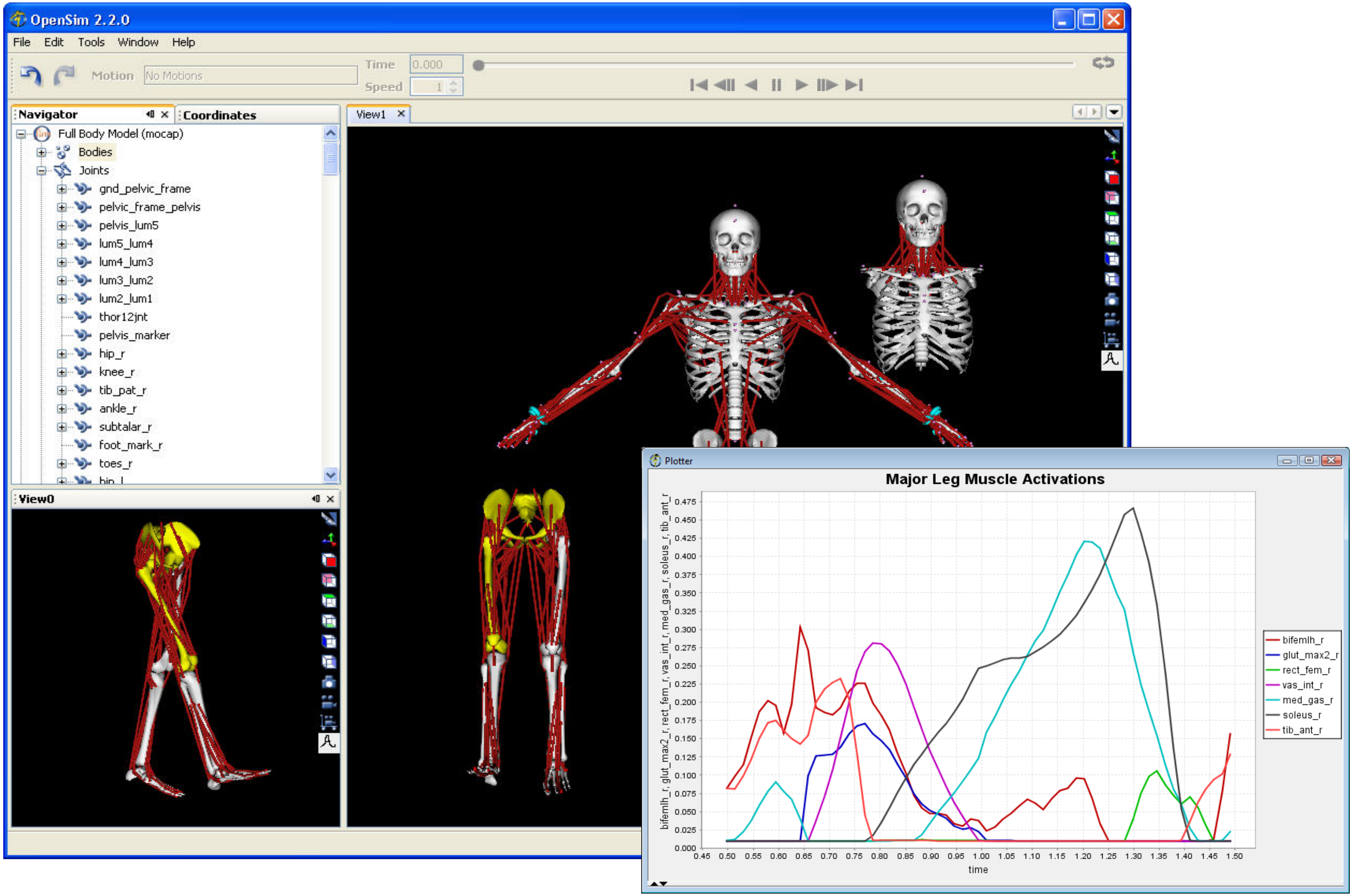
Visualize human running in detail



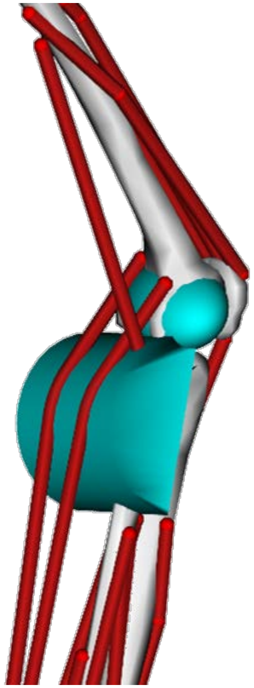
Examine causes ankle injury



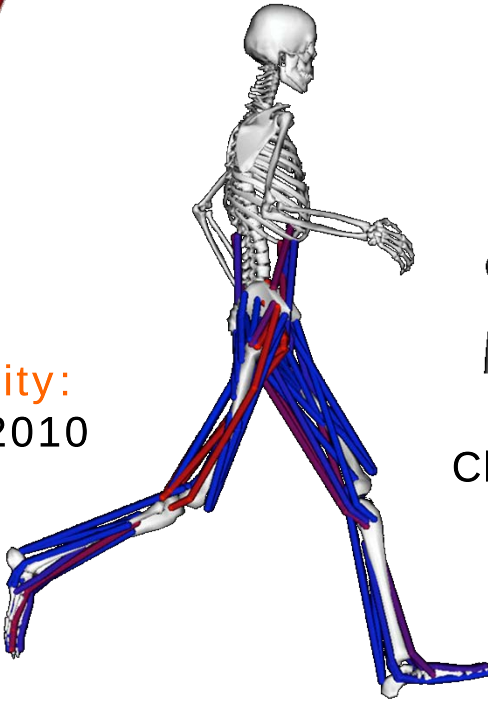
OpenSim is an application



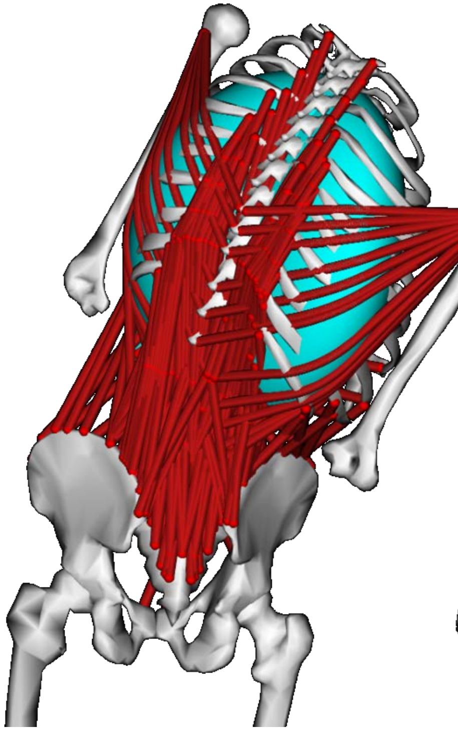
OpenSim is a repository of models, data, & tools



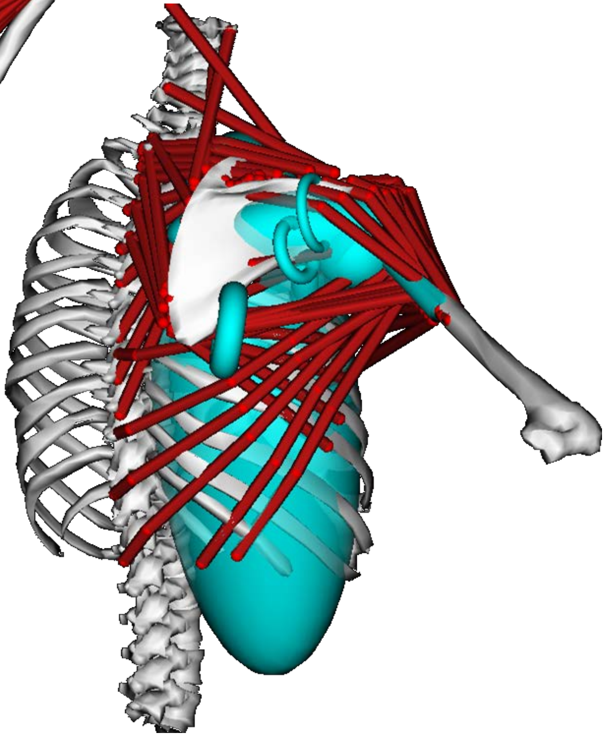
Lower-extremity:
Arnold et al, 2010



Running: Hamner et al, 2010

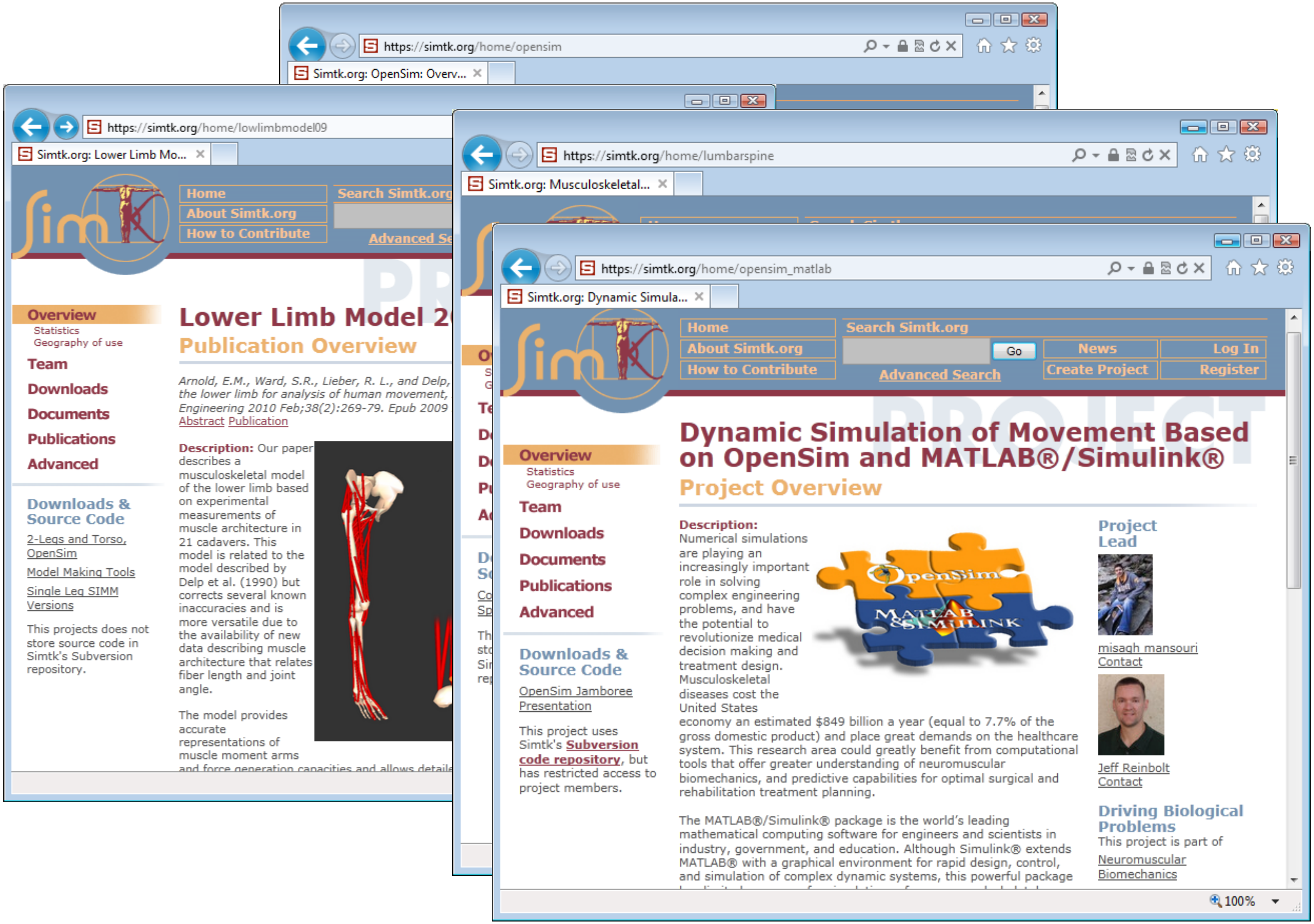


Lumbar-spine:
Christophy et al, 2011



Shoulder:
Matias et al, in prep.

OpenSim is a repository of models, data, & tools



https://simtk.org/home/opensim

Simtk.org: OpenSim: Overv...

https://simtk.org/home/lowlimbmodel09

Simtk.org: Lower Limb Mo...

This projects does not store source code in Simtk's Subversion repository.

Lower Limb Model 2

Publication Overview

Arnold, E.M., Ward, S.R., Lieber, R. L., and Delp, the lower limb for analysis of human movement, Engineering 2010 Feb;38(2):269-79. Epub 2009
Abstract Publication

Description: Our paper describes a musculoskeletal model of the lower limb based on experimental measurements of muscle architecture in 21 cadavers. This model is related to the model described by Delp et al. (1990) but corrects several known inaccuracies and is more versatile due to the availability of new data describing muscle architecture that relates fiber length and joint angle.

The model provides accurate representations of muscle moment arms and force generation capacities and allows detail

https://simtk.org/home/lumbarspine

Simtk.org: Musculoskeletal...

https://simtk.org/home/opensim_matlab

Simtk.org: Dynamic Simula...

This project uses Simtk's **Subversion code repository**, but has restricted access to project members.

Dynamic Simulation of Movement Based on OpenSim and MATLAB®/Simulink®

Project Overview

Description: Numerical simulations are playing an increasingly important role in solving complex engineering problems, and have the potential to revolutionize medical decision making and treatment design. Musculoskeletal diseases cost the United States economy an estimated $849 billion a year (equal to 7.7% of the gross domestic product) and place great demands on the healthcare system. This research area could greatly benefit from computational tools that offer greater understanding of neuromuscular biomechanics, and predictive capabilities for optimal surgical and rehabilitation treatment planning.

The MATLAB®/Simulink® package is the world's leading mathematical computing software for engineers and scientists in industry, government, and education. Although Simulink® extends MATLAB® with a graphical environment for rapid design, control, and simulation of complex dynamic systems, this powerful package

Project Lead

misagh mansouri
Contact

Jeff Reinbolt
Contact

Driving Biological Problems

This project is part of Neuromuscular Biomechanics

13,000 users of these tools; 70 projects with downloads

OpenSim is a resource

http://opensim.stanford.edu

NCSRR
NATIONAL CENTER
FOR SIMULATION IN
REHABILITATION
RESEARCH

OpenSim Community

- SEE THE WORK
- JOIN THE COMMUNITY
- FIND SUPPORT, EVENTS, & RESOURCES

OpenSim

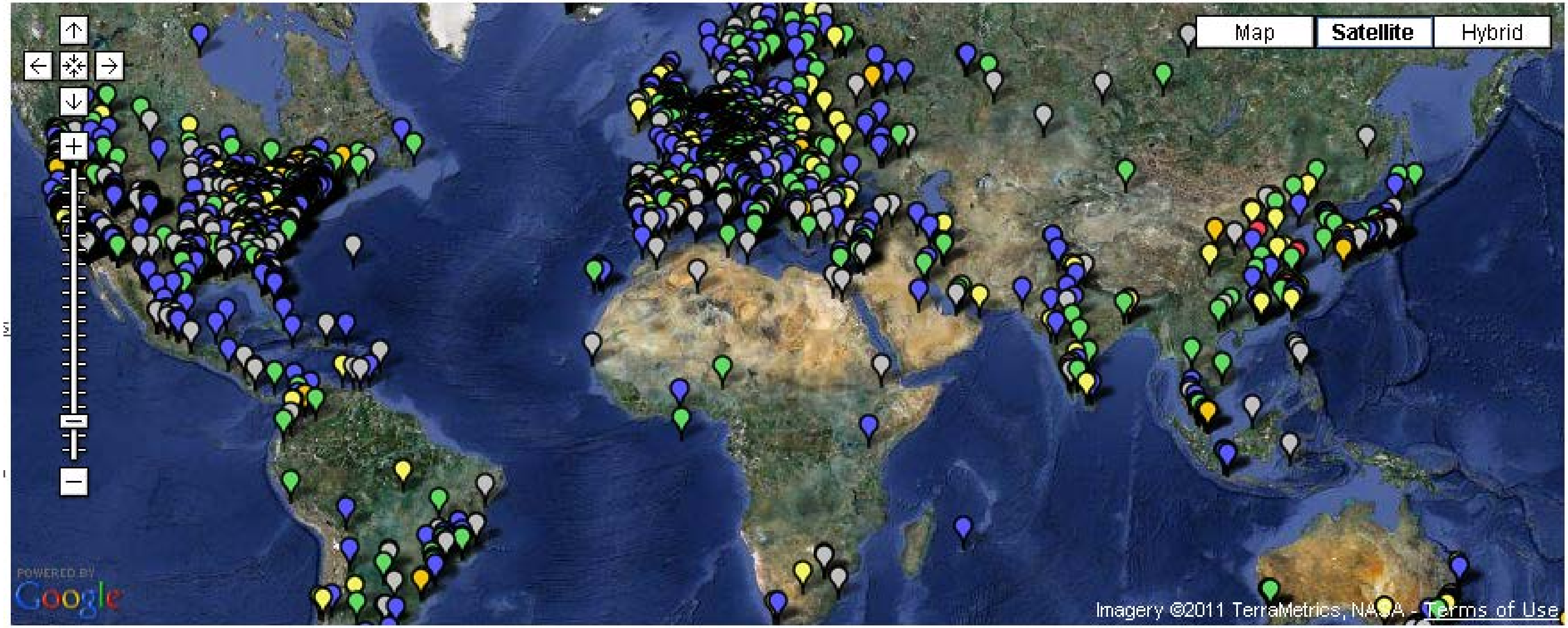
State-of-the-art
simulation software
advancing research
in rehabilitation science

SEE THE GREAT WORK,
JOIN THE OPENSIM COMMUNITY
TO GET STARTED, AND
FIND THE SUPPORT, EVENTS,
& RESOURCES YOU NEED
TO SUCCEED.

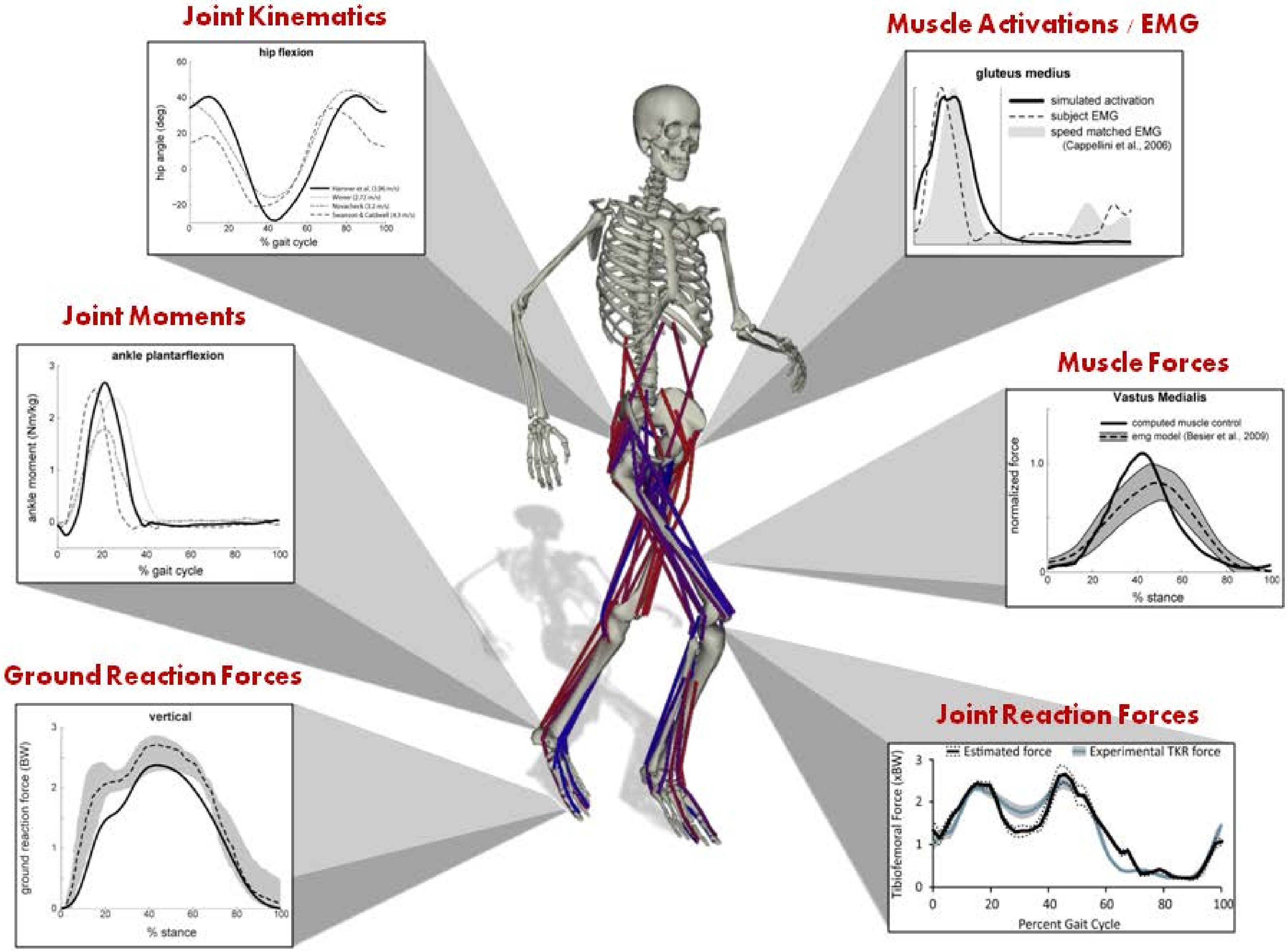
OpenSim is a worldwide community

86702 Page Hits in the past 180 Days (9742 Unique Visitors)

2345 Stanford Page Hits (81 Unique Visitors)



Testing the Accuracy of Simulations



Joint Kinematics

Muscle Activations / EMG

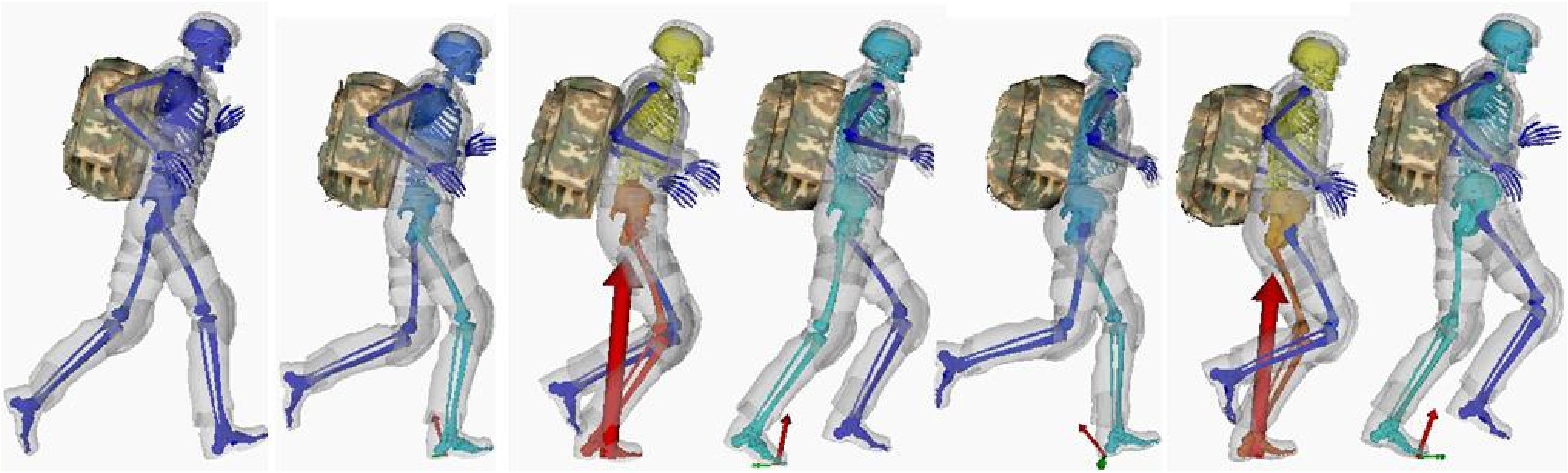
Joint Moments

Muscle Forces

Ground Reaction Forces

Joint Reaction Forces

Motion Synthesis

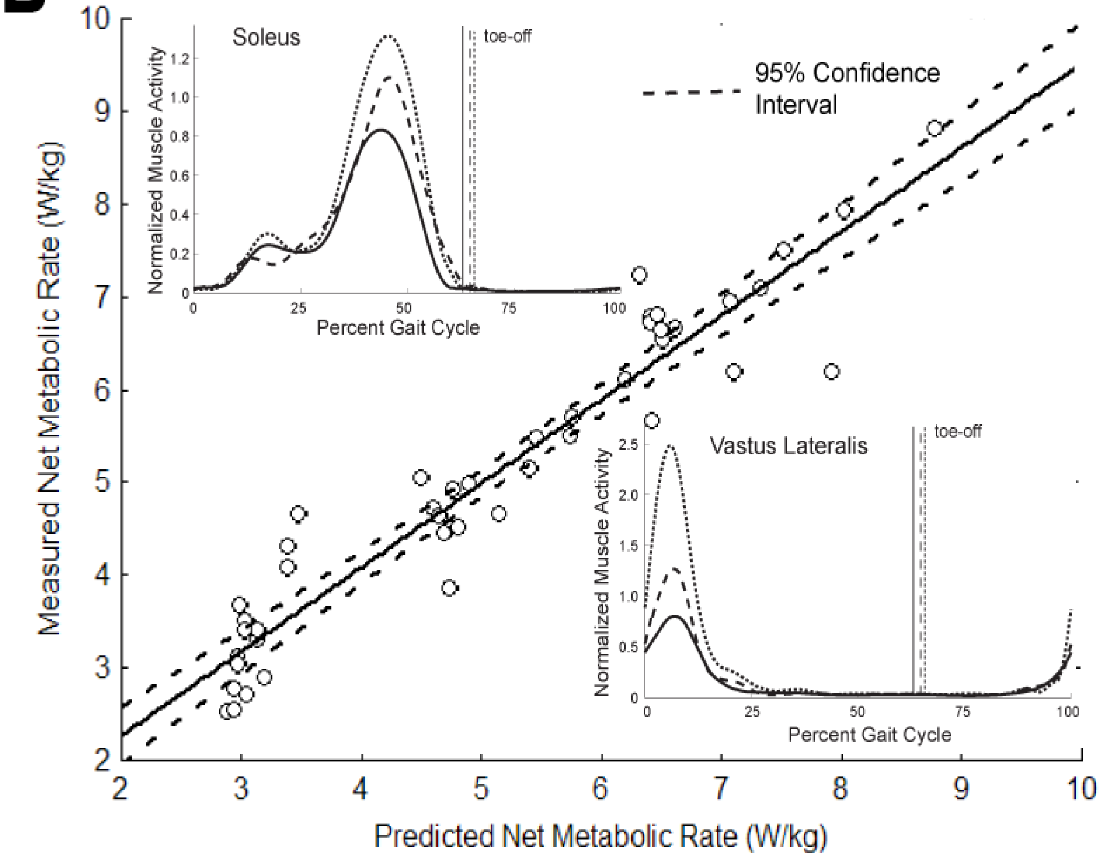


Wang, JM, Hamner, SR, Delp, SL, Koltun, V. Optimizing locomotion controllers using biologically-based actuators and objectives. *SIGGRAPH 2012*

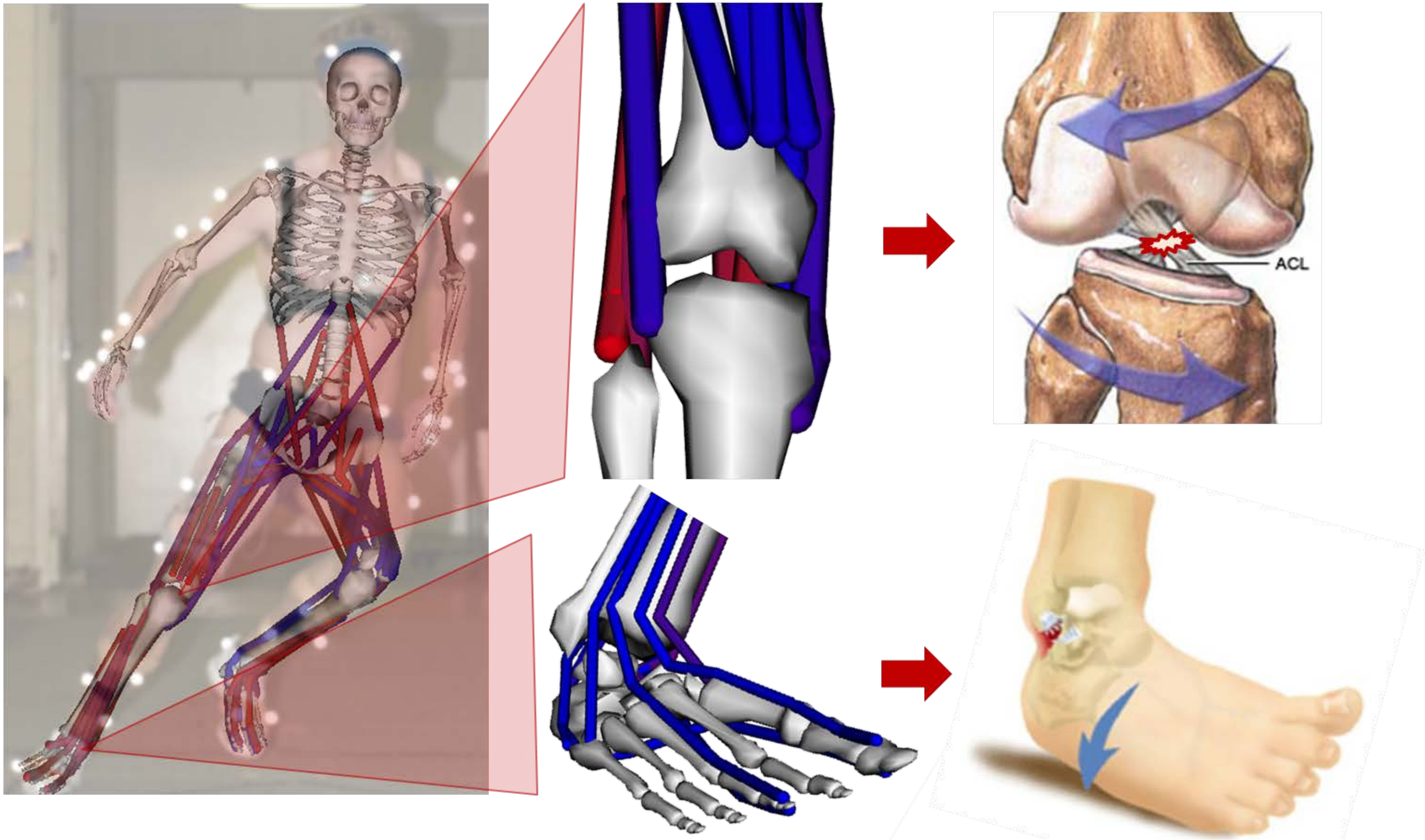
Metabolics Assessment

A

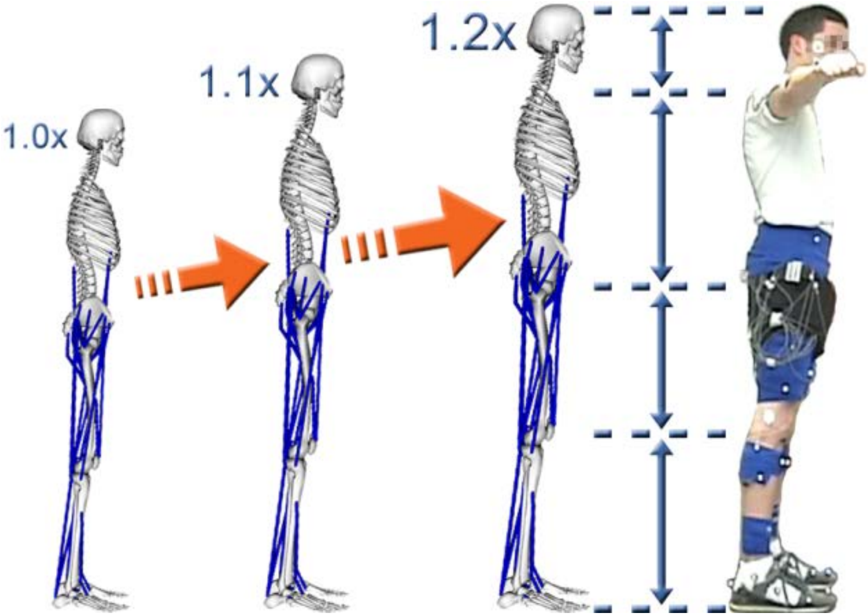
B

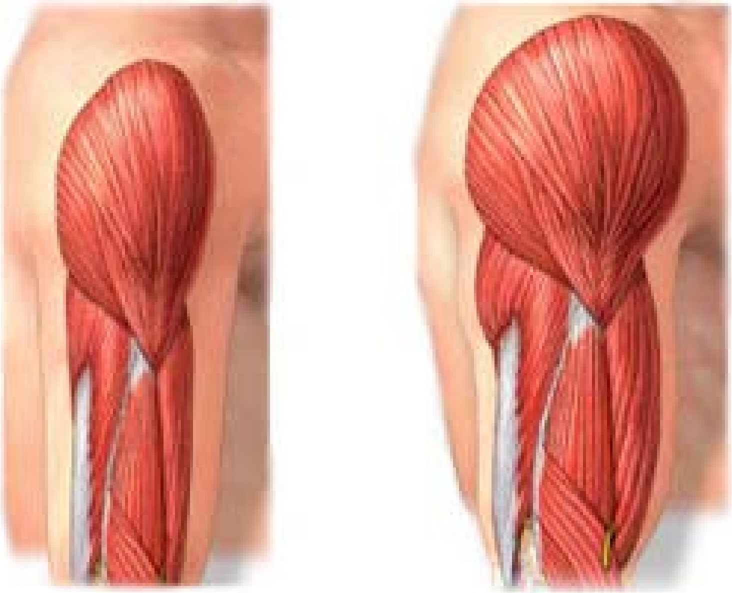


Injury Risk Assessment



Soldier-Specific Scaling





Support, Training, and Collaboration to Prove Your Technology

OpenSim

Project Overview

Description: OpenSim is a freely available, user extensible software system that lets users develop models of musculoskeletal structures and create dynamic simulations of movement.

Find out how to join the OpenSim community at opensim.stanford.edu.

The software provides a platform on which the biomechanics community can build a library of simulations that can be exchanged, tested, analyzed, and improved through multi-institutional collaboration. The underlying software is written in ANSI C++, and the graphical user interface (GUI) is written in Java.

OpenSim technology makes it possible to develop customized controllers, analyses, contact models, and muscle models among other things. These plugins can be shared without the need to alter or compile source code. Users can analyze existing models and simulations and develop new models and simulations from within the GUI.

Watch the video below to get an overview of the OpenSim project and see how modeling can be used to help plan surgery for children with cerebral palsy:
http://www.youtube.com/watch?v=MEOVHfCtIM0

Downloads & Source Code

OpenSim
OpenSim Release Betas
OpenSim Workshops

This project uses Simtk's **Subversion code repository**, but has restricted access to project members.

News

OpenSim 2.4 Released
OpenSim at the ISB Congress
OpenSim 2.2.1 Released

Project Lead

Scott Delp
Contact

Ayman Habib
Contact

Ajay Seth
Contact

Jennifer Hicks
Contact
[Show Fewer Leads]

Load OpenSim on your laptop for tomorrow’s hands-on workshop

Training the next generation

OpenSim is a team of contributors:


Scott Delp


Ayman Habib


Jennifer Hicks


Jeff Reinbolt


Ajay Seth


Michael Sherman


Edith Arnold


Matt DeMers


Sam Hamner


Chand John


Kat Steele


Melanie Fox


Peter Eastman


Clay Anderson


Allison Arnold


Eran Guendelman

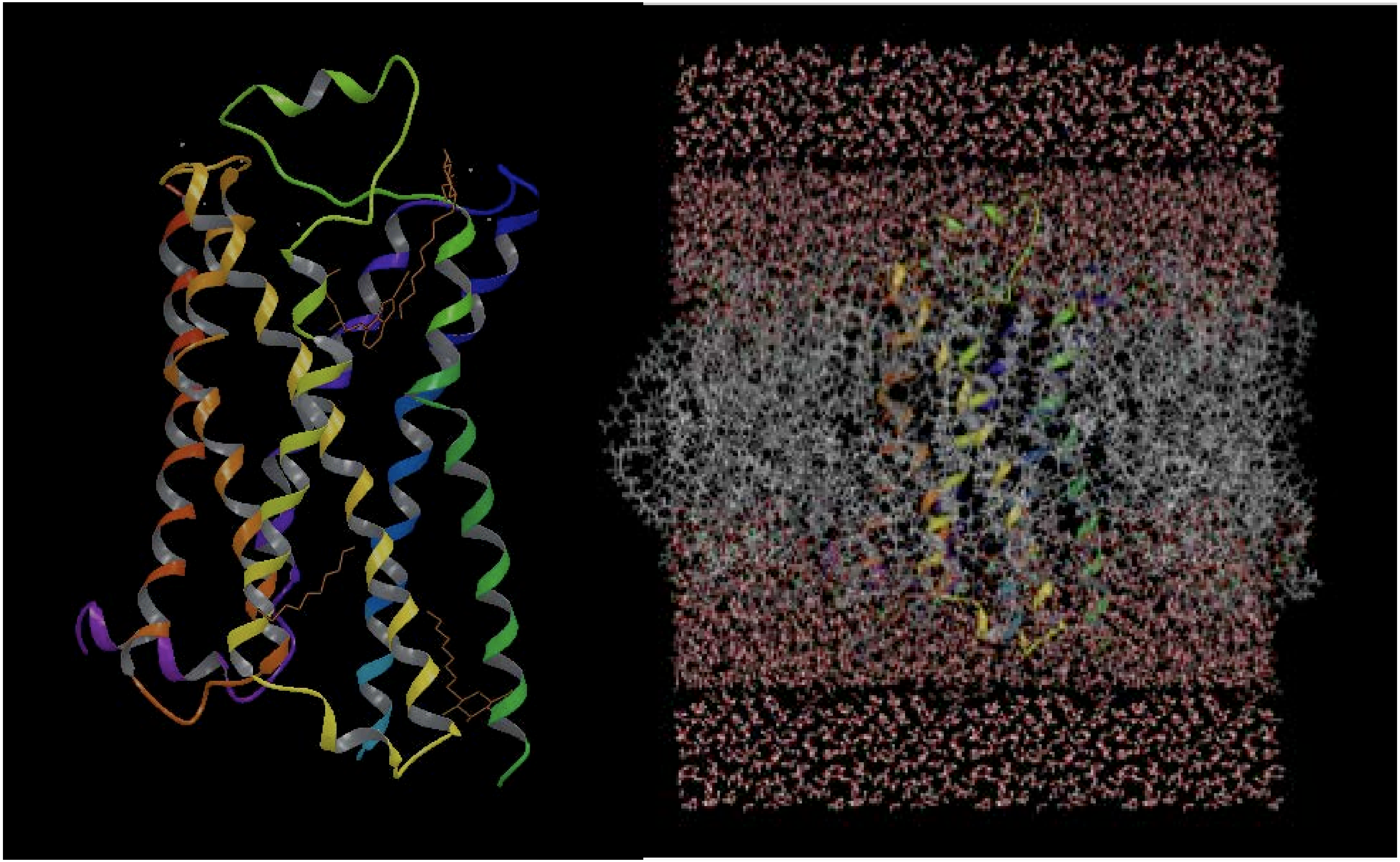

May Liu


Peter Loan


Darryl Thelen

WW Teams

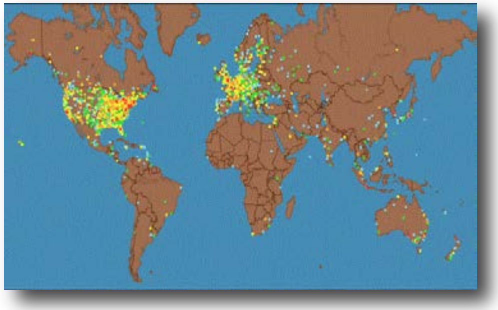
Molecular Dynamics Simulations: faster, bigger, easier to use and apply



Our solution to the long-timescale simulation challenge

Combine multiple, powerful, complementary technologies

1) Folding@home: very large-scale distributed computing



Most powerful computer cluster in the world (~8 petaflops)

10^4x to 10^5x

http://folding.stanford.edu

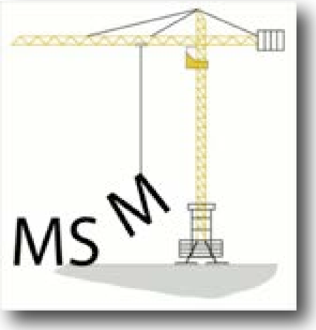
Voelz, *et al*, *JACS* (2010)
Ensign *et al*, *JMB (2007)*
Shirts and Pande, *Science* (2000)

2) OpenMM: Very fast MD (~1µs/day) on GPUs

~1µs/day for implicit solvent simulaton of small proteins (~40aa)

~10^2x

3) Markov State Models (MSMs): Stat mech of many trajectories



very long timescale dynamics by combining many simulations

10^2x to 10^3x

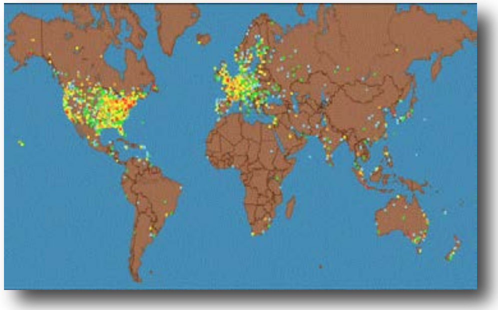
http://simtk.org/home/msmbuilder

Bowman, *et al*, *J. Chem. Phys.* (2009)
Singhal & Pande, *J. Chem. Phys.* (2005)

Our solution to the long-timescale simulation challenge

Combine multiple, powerful, complementary technologies

1) Folding@home: very large-scale distributed computing



Most powerful computer cluster in the world (~8 petaflops)

10^4x to 10^5x

http://folding.stanford.edu

Voelz, *et al*, *JACS* (2010)
Ensign *et al*, *JMB (2007)*
Shirts and Pande, *Science* (2000)

2) OpenMM: Very fast MD (~1µs/day) on GPUs

~1µs/day for implicit solvent simulaton of small proteins (~40aa)

~10^2x

http://simtk.org/home/openmm

Elsen, *et al. ACM/IEEE conf. on Supercomputing* (2006)
Friedrichs, *et al. J. Comp. Chem.*, (2009)
Eastman and Pande. *J. Comp. Chem.* (2009)

3) Markov State Models (MSMs): Stat mech of many trajectories

very long timescale dynamics by combining many simulations

10^2x to 10^3x

Phys. (2009)

OpenMM: http://simtk.org/home/openmm

- **What is OpenMM?**
  - API, library, and application for core molecular dynamics
  - open source (**LGPL**) software
- **Key features**
  - Very rapid execution (GPUs, multi-core CPUs)
  - rapid development: clean API, python bindings, custom force classes
  - support of modern hardware
  - broad support for forcefields with full acceleration:
    - additive: AMBER, CHARMM, OPLS, ROSETTA
    - polarizable/more accurate: AMOEBA, RPMD
    - implicit solvent: GB/SA variants, EEF1, ABSINTH
- **OpenMM enabled applications**
  - GROMACS, Folding@home, Protomol, AMBER (file formats), CHARMM (C. Brooks)

http://simtk.org/home/openmm

OpenMM: Fast execution *and rapid development*

- **Interface to Python: towards a Domain Specific Language (DSL)**
  - 10 lines of code to a customizable, high performance MD code
  - tweak to your hearts content, but keep high performance aspects

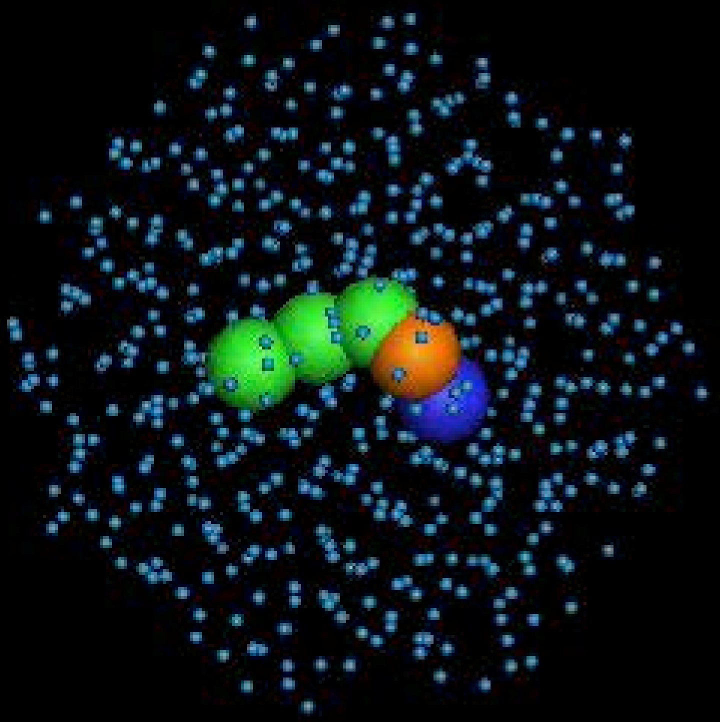
```
from simtk.openmm.app import *
from simtk.openmm import *
from simtk.unit import *

pdb = PDBFile('input.pdb')
forcefield = ForceField('amber99sb.xml', 'tip3p.xml')
system = forcefield.createSystem(pdb.topology, nonbondedMethod=PME,
    nonbondedCutoff=1*nanometer, constraints=HBonds)
integrator = LangevinIntegrator(300*kelvin, 1/picosecond, 0.002*picoseconds)
simulation = Simulation(pdb.topology, system, integrator)
simulation.context.setPositions(pdb.positions)
simulation.minimizeEnergy()
simulation.reporters.append(PDBReporter('output.pdb', 1000))
simulation.step(10000)
```

- **Custom Force classes**
  - code in equations, rather than CUDA or OpenCL, but keeping high performance
  - Custom classes for implicit solvent, bonded, non-bonded, integrators
  - looking to generalize further based on interest from the community

```
force = CustomBondForce("D*(1-exp(a*(r0-r)))^2")
```

Application to biomolecular assembly

(Wagoner, VSP)

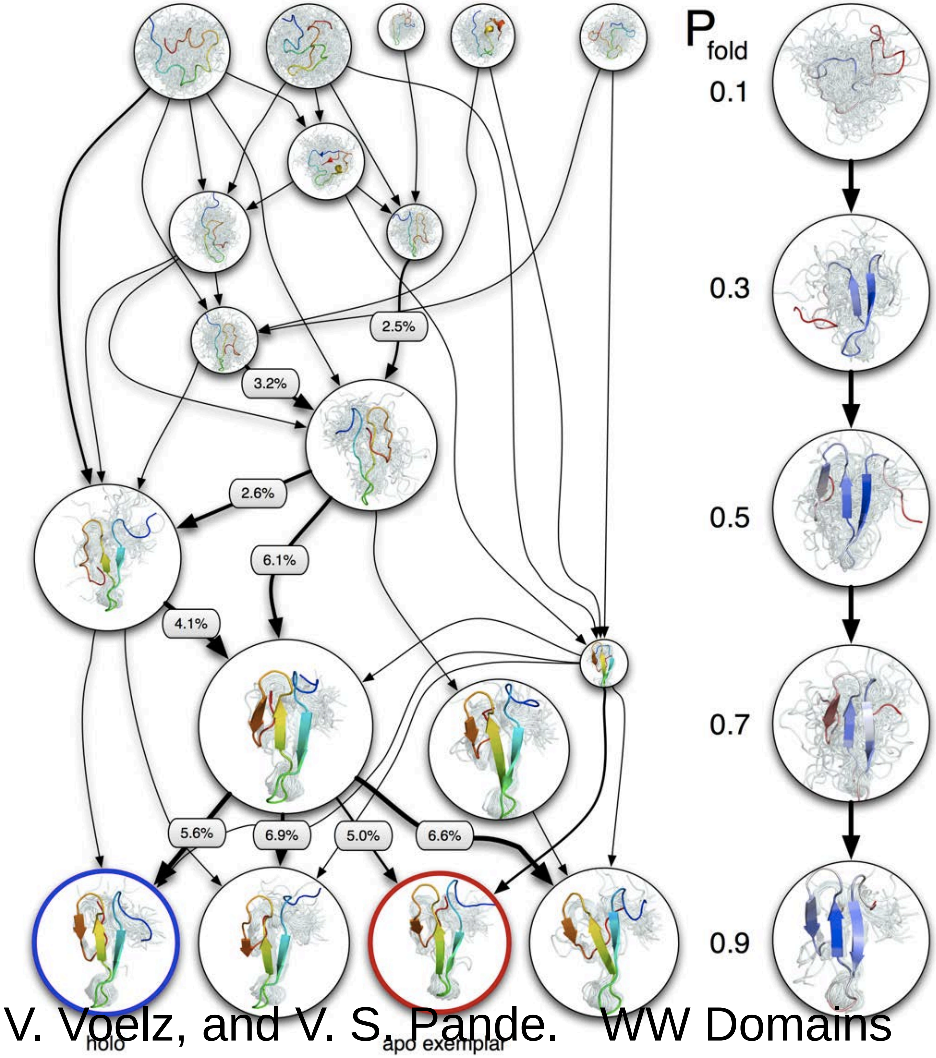


Grand Canonical construction of a micelle

Movie made with VMD

MSMs yield novel insight:
Nature of folding pathways

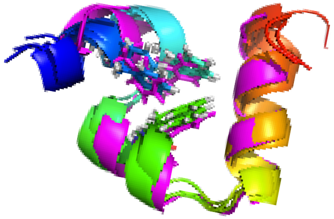
- **Simply having long trajectories does not solve the challenge of analysis**
  - how to gain insight from trajectories?
- **MSM finds heterogeneous pathways, not a single path**
  - simple analyses of long trajectory data did not find
  - original method assumed single pathway
  - MSMs don't make these assumptions (eg WW simulations from Noé et al, *PNAS* 2009)



T. J. Lane, G. Bowman, K. Beauchamp, V. Voelz, and V. S. Pande. WW Domains Take a New Turn: Markov State Model Reveals Folding and Functional Dynamics in Ultra-Long MD Trajectories. *JACS* (2011)

General Result #1: This approach is broadly applicable

Summary of our key protein folding simulation successes

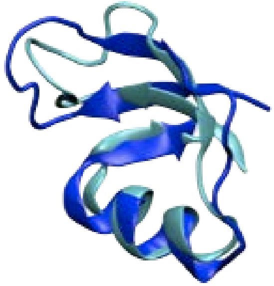


villin

36 amino acids

microsecond folding time

published in: 2001, 2007

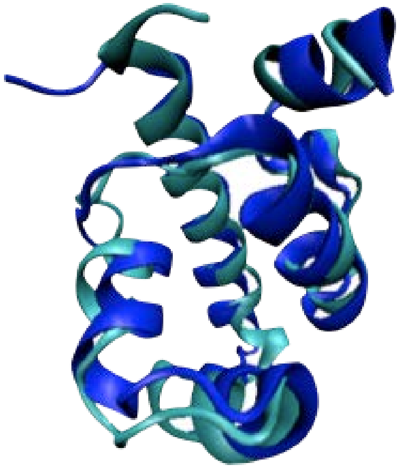


NTL9

39 amino acids

millisecond folding time

published in:2009

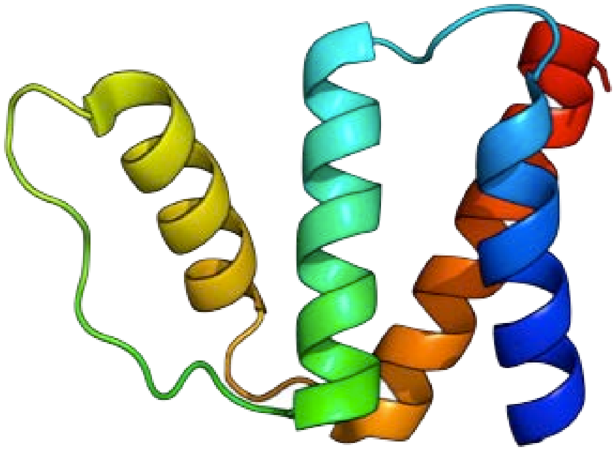


λ-repressor

80 amino acids

millisecond folding time

published in: 2010



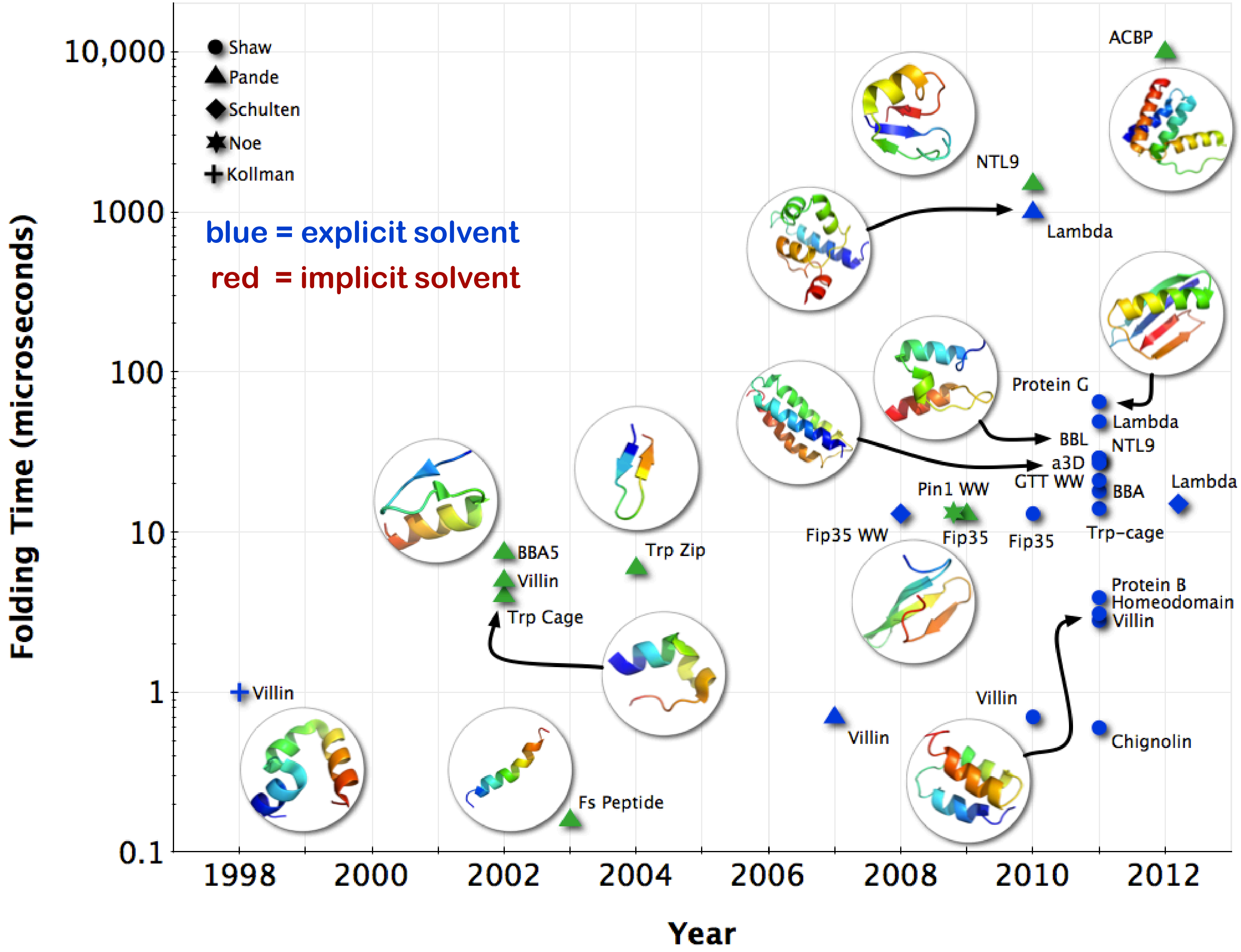
ACBP

85 amino acids

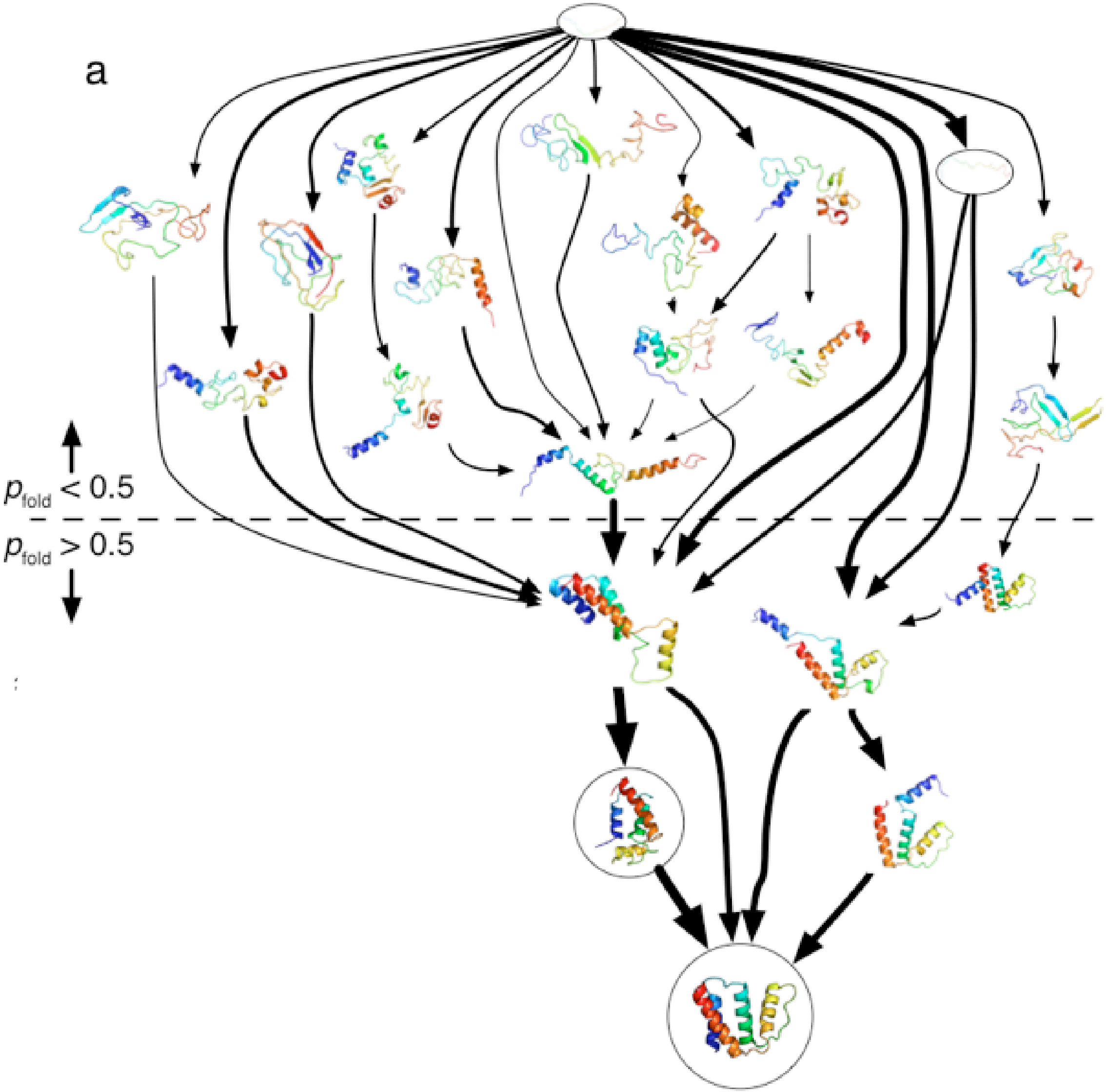
10 millisecond folding time

submitted in: 2011

What has the simulation community done so far?

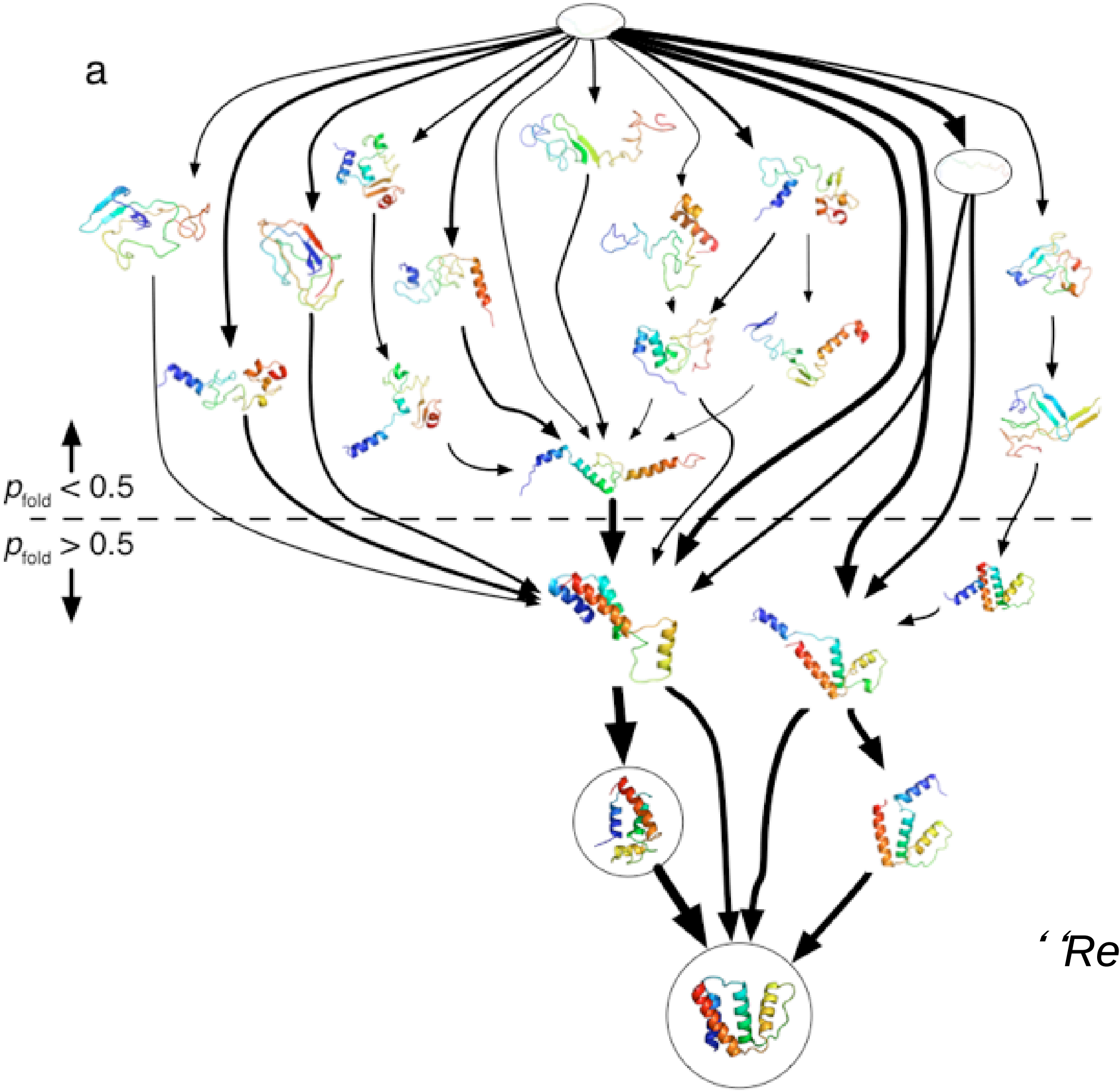


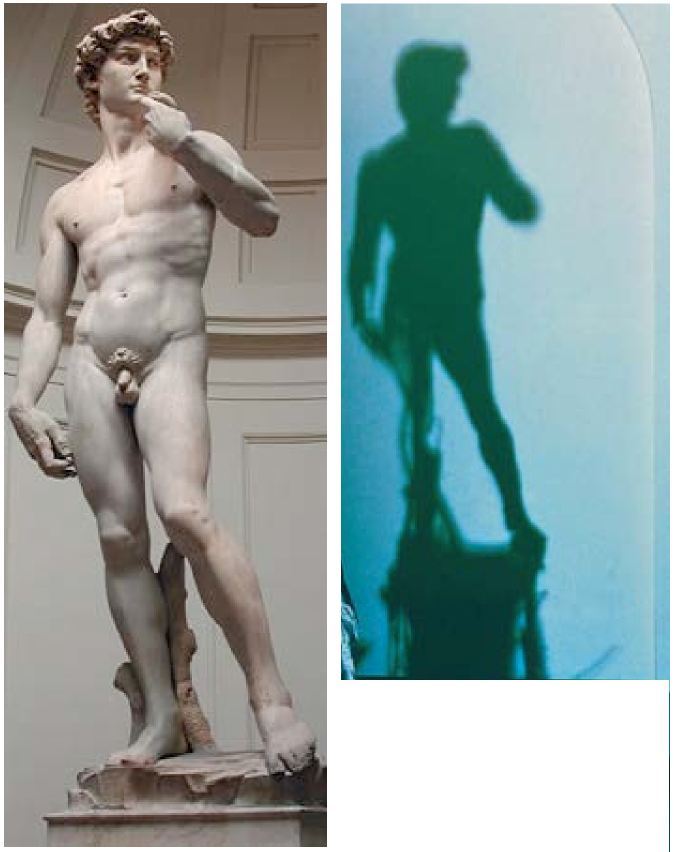
General result #2: Kinetics looks like transitions between many metastable states: *ACBP example below*



General result #3: Consequences of projections

How can one reconcile this with the simple picture?



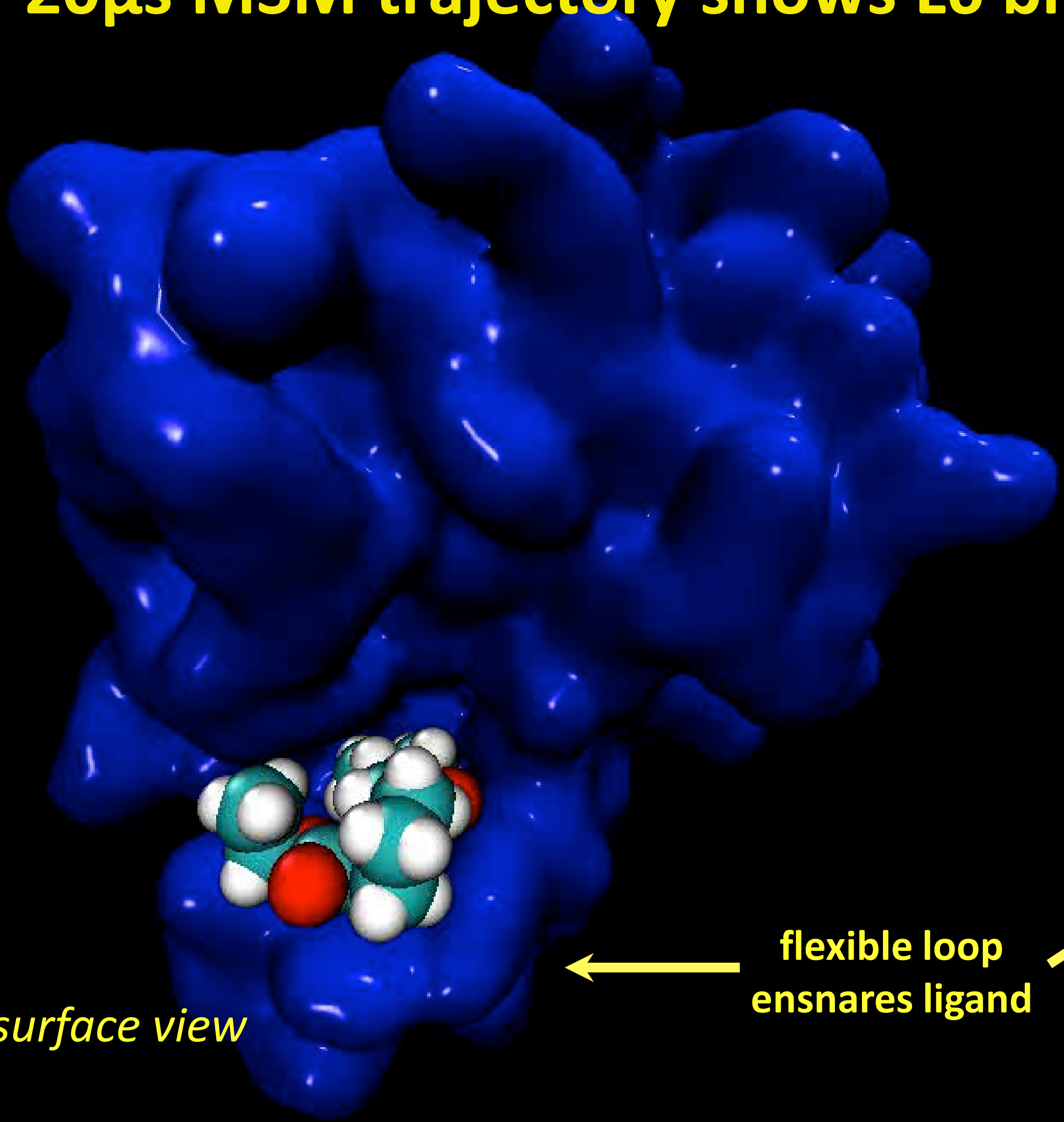


‘‘Regarded from two sides’’ by Diet Wiegman (1984)
Kruschela & Zagrovic.
DOI:10.1039/b917186j

(Lawrenz, VSP)

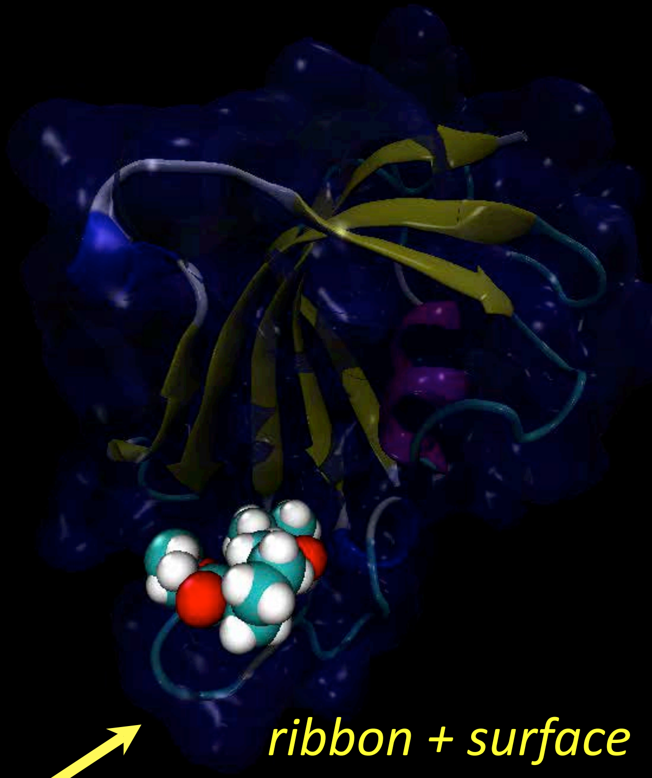
movie made with VMD

20µs MSM trajectory shows L6 binding



surface view

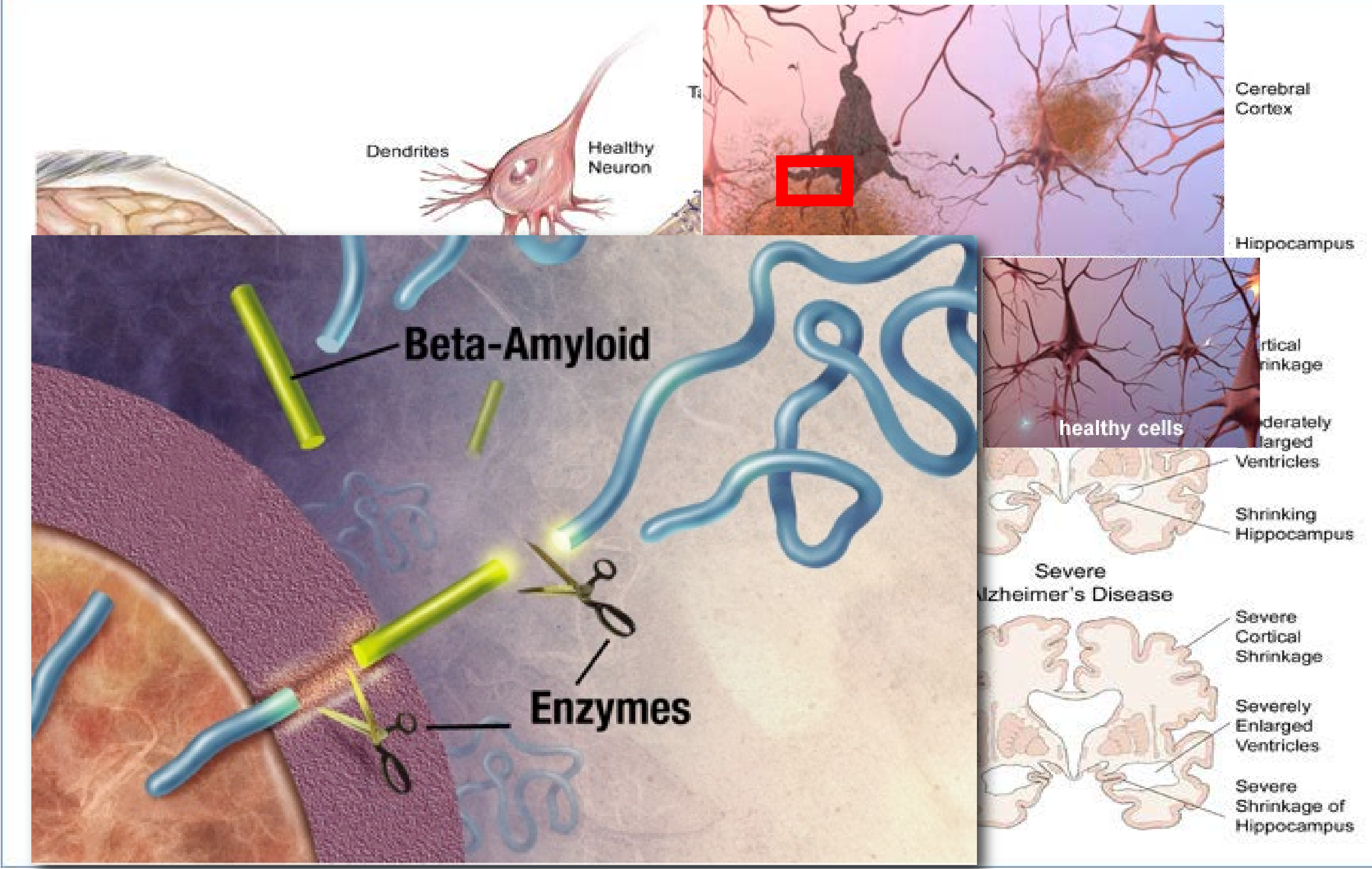
flexible loop ensnares ligand



ribbon + surface view

Movie made with VMD

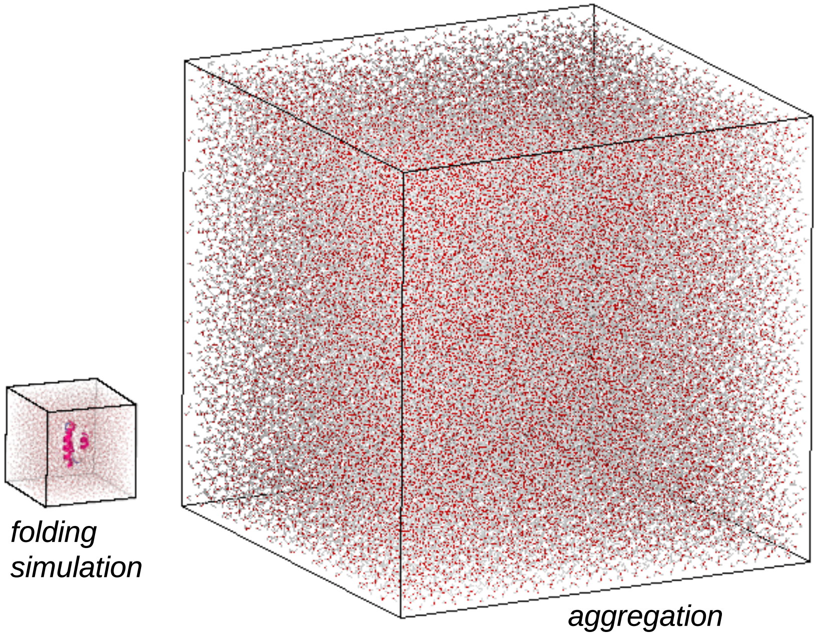
Application of MSMs to misfolding:
Simulation of misfolding in Alzheimer's Disease



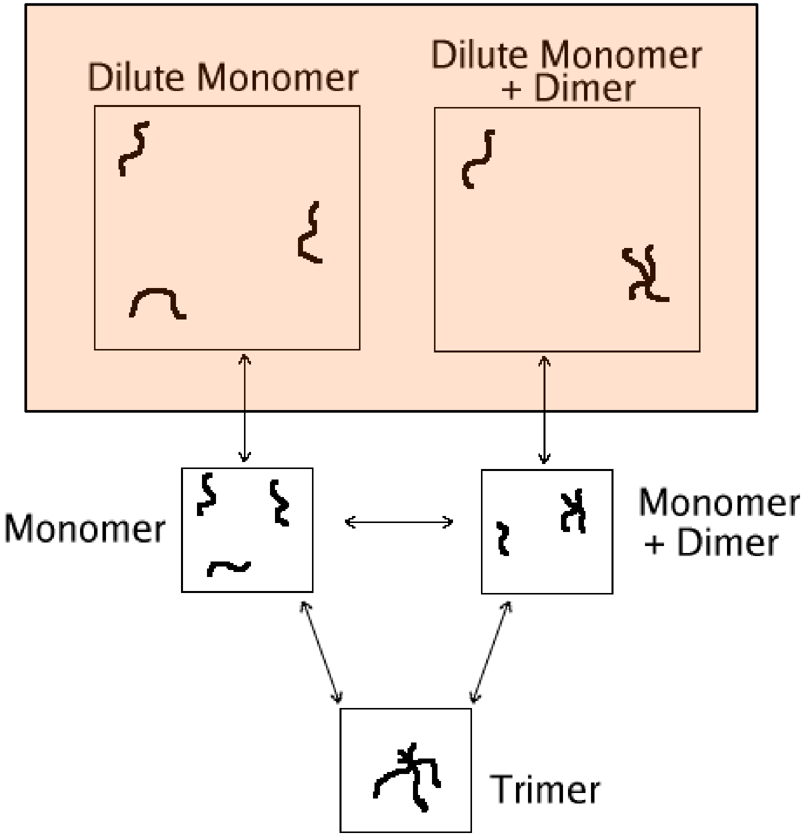
Extensions to MSMs for oligomerization

Large system size needed in aggregation simulation:

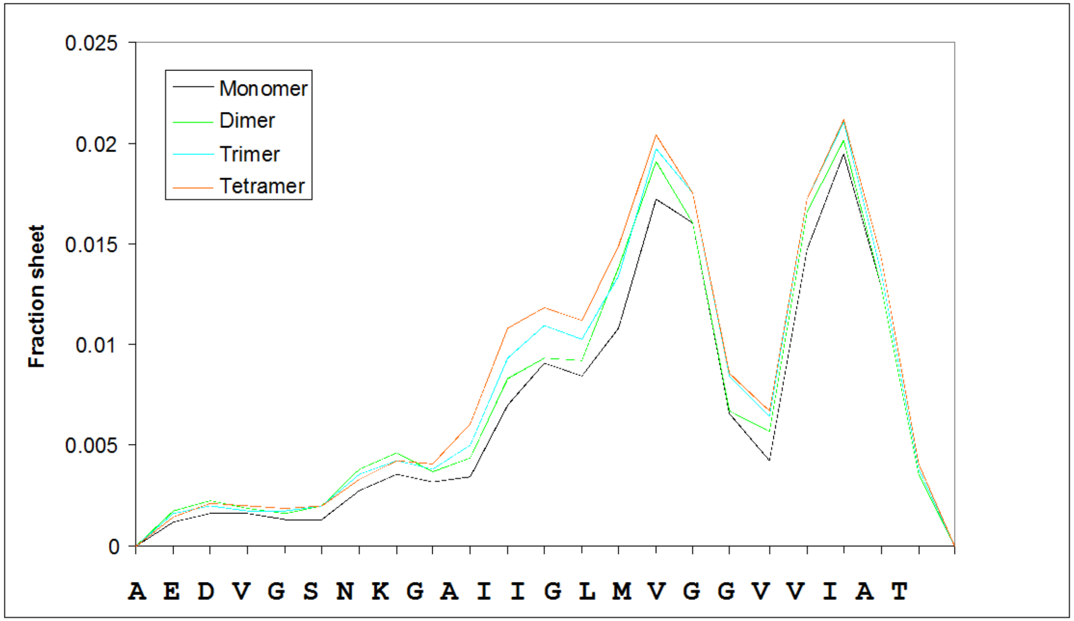
Folding simulation compared to aggregation simulation

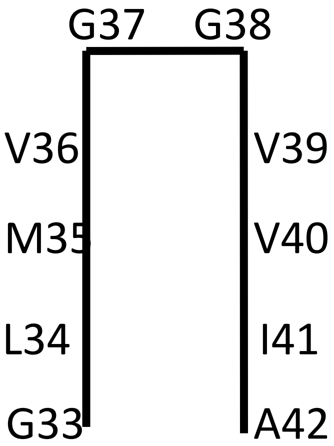


Adding low concentration states allows for a direct comparison to experiment



We find a ß-hairpin structure





G37 G38

V36 V39

M35 V40

L34 I41

G33 A42

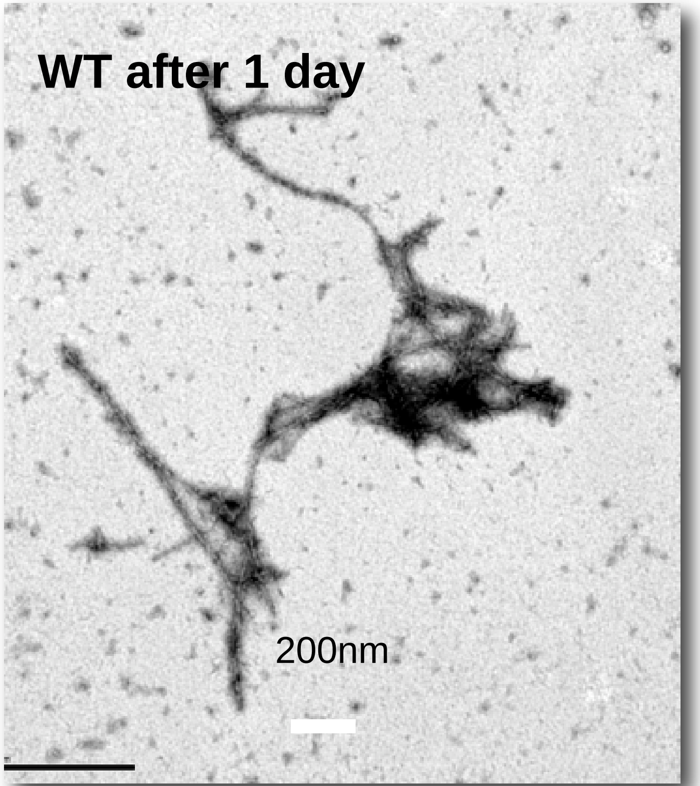
hairpin greatly stabilized by 41-42

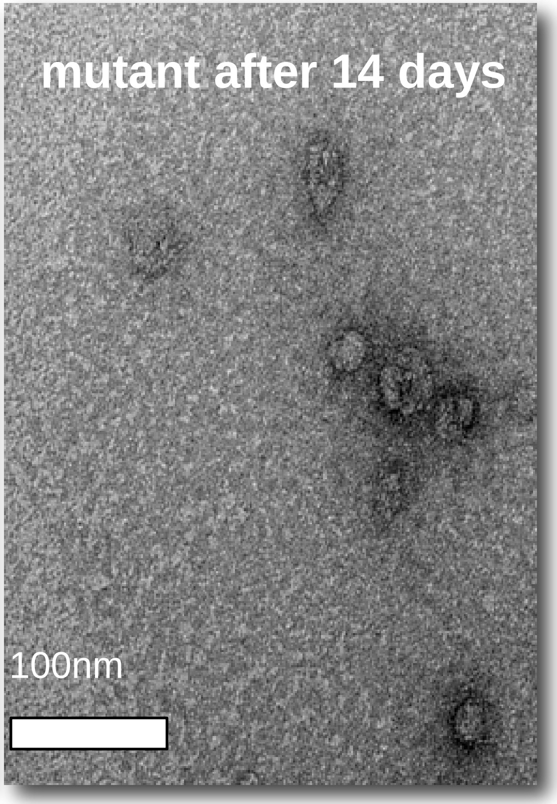
- We find a hairpin centered around residues 37-38
- let’s stabilize the hairpin to force the structure we see
- G37D-pro mutant to fix hairpin structure
- **yields a theory to explain why 1-42 is different from 1-40**

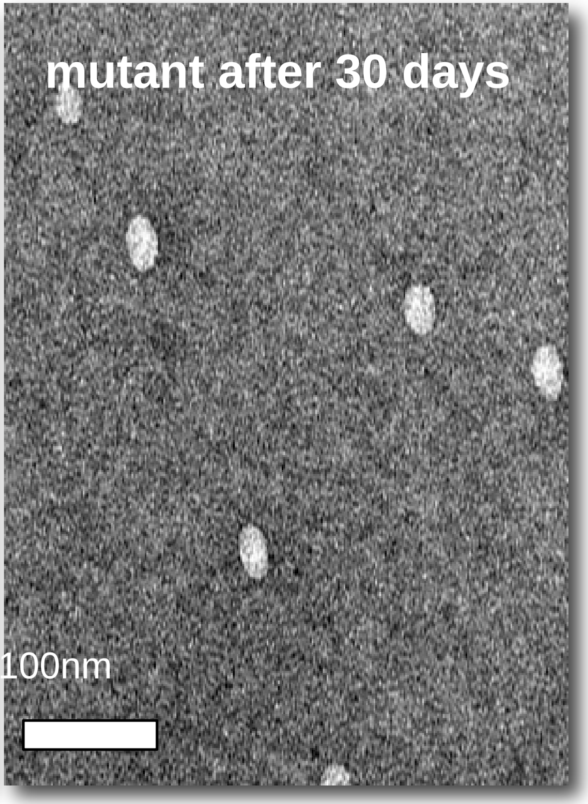
V. Vishal, N. W. Kelley, R. Jaykumar, C. Liu, T. Pray, G. Krafft, and V. S. Pande. Simulation of the rate of formation and structure of Aß oligomers. Journal of Chemical Physics, 129 214707 (2008)

(Liu, Jayakumar, VSP)

TEM: Mutant shows slowed aging







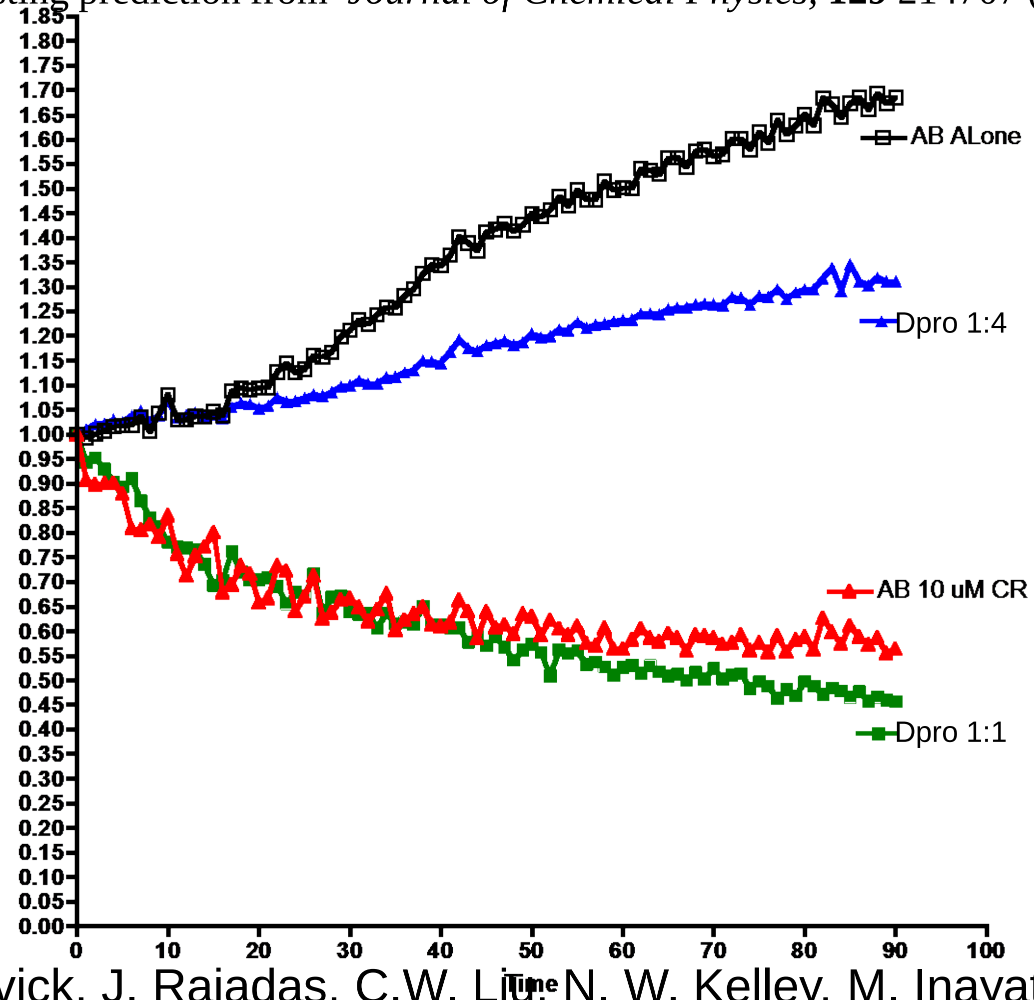
testing prediction from *Journal of Chemical Physics*, **129** 214707 (2008)

- **WT forms protofibrils easily (days)**
  - Mutant considerably more stable
  - We see spheriod particles, but no fibrils

P. Novick, J. Rajadas, C.W. Liu, N. W. Kelley, M. Inayathullah, and V. S. Pande. Rationally designed turn promoting mutation in the amyloid β-peptide sequence to stabilize oligomers in solution. *PLoS ONE* **6**(7): e21776 (2011)

Dpro mutant inhibits WT aggregation

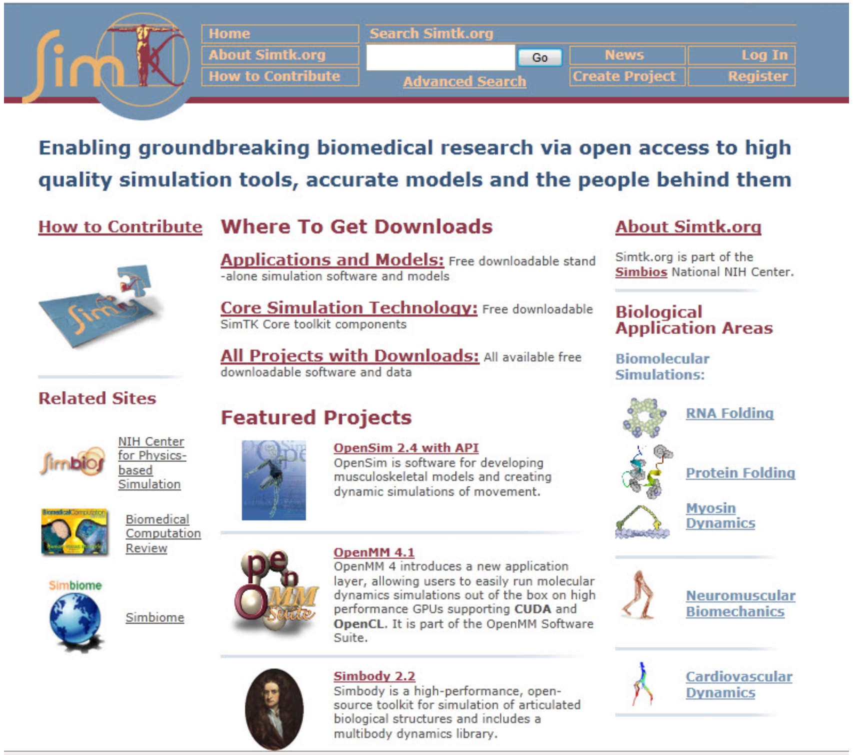
testing prediction from *Journal of Chemical Physics*, **129** 214707 (2008)



- **Mutant peptide inhibits WT aggregation**
  - great reduction in aggregation compared to WT
- **Results comparable to known small molecule aggregation inhibitors (eg congo red)**
- **Suggests new small molecules**

P. Novick, J. Rajadas, C.W. Liu, N. W. Kelley, M. Inayathullah, and V. S. Pande. Rationally designed turn promoting mutation in the amyloid β-peptide sequence to stabilize oligomers in solution. *PLoS ONE* **6**(7): e21776 (2011)

Simtk.org: a resource for sharing code, data, models



Enabling groundbreaking biomedical research via open access to high quality simulation tools, accurate models and the people behind them

How to Contribute

Related Sites

NIH Center for Physics-based Simulation

Biomedical Computation Review

Simbiome

Where To Get Downloads

Applications and Models: Free downloadable stand-alone simulation software and models

Core Simulation Technology: Free downloadable SimTK Core toolkit components

All Projects with Downloads: All available free downloadable software and data

Featured Projects

OpenSim 2.4 with API
OpenSim is software for developing musculoskeletal models and creating dynamic simulations of movement.

OpenMM 4.1
OpenMM 4 introduces a new application layer, allowing users to easily run molecular dynamics simulations out of the box on high performance GPUs supporting **CUDA** and **OpenCL**. It is part of the OpenMM Software Suite.

Simbody 2.2
Simbody is a high-performance, open-source toolkit for simulation of articulated biological structures and includes a multibody dynamics library.

About Simtk.org

Simtk.org is part of the Simbios National NIH Center.

Biological Application Areas

Biomolecular Simulations:

RNA Folding

Protein Folding

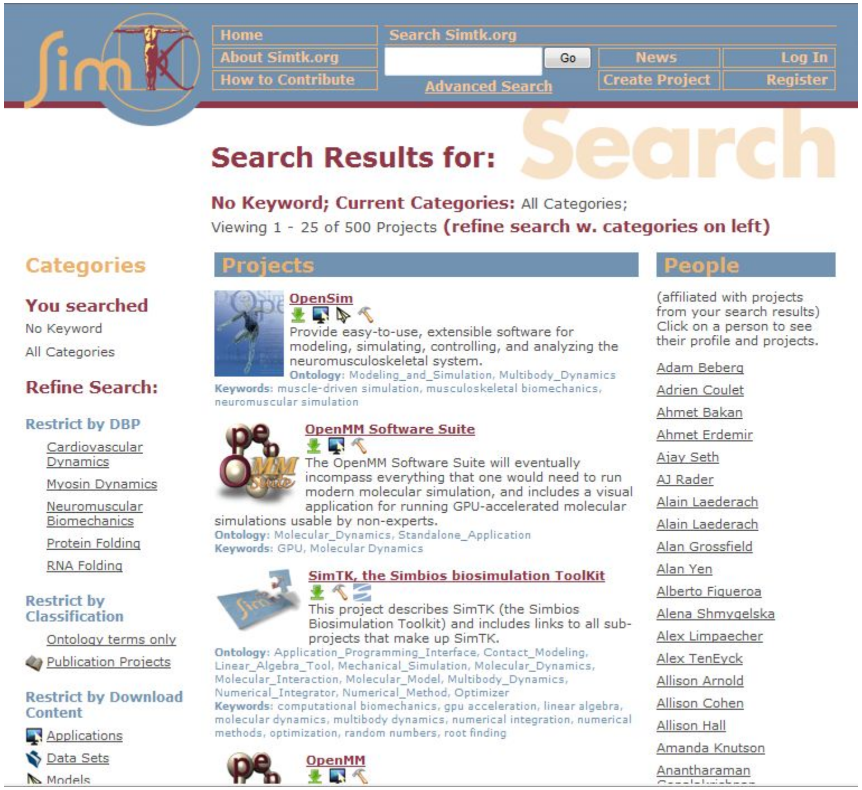
Myosin Dynamics

Neuromuscular Biomechanics

Cardiovascular Dynamics

- Web-based project hosting site for the biocomputational community
- Features include:
  - Privacy options to share data/software only when you are ready
  - Daily backups
  - Statistics on project visitors and users

Simtk.org Projects



Search Results for:

No Keyword; Current Categories: All Categories;
Viewing 1 - 25 of 500 Projects **(refine search w. categories on left)**

Categories

You searched
No Keyword
All Categories

Refine Search:

Restrict by DBP
- Cardiovascular Dynamics
- Myosin Dynamics
- Neuromuscular Biomechanics
- Protein Folding
- RNA Folding

Restrict by Classification
- Ontology terms only
- Publication Projects

Restrict by Download Content
- Applications
- Data Sets
- Models

Projects

OpenSim
Provide easy-to-use, extensible software for modeling, simulating, controlling, and analyzing the neuromusculoskeletal system.
Ontology: Modeling_and_Simulation, Multibody_Dynamics
Keywords: muscle-driven simulation, musculoskeletal biomechanics, neuromuscular simulation

OpenMM Software Suite
The OpenMM Software Suite will eventually incompass everything that one would need to run modern molecular simulation, and includes a visual application for running GPU-accelerated molecular simulations usable by non-experts.
Ontology: Molecular_Dynamics, Standalone_Application
Keywords: GPU, Molecular Dynamics

SimTK, the Simbios biosimulation ToolKit
This project describes SimTK (the Simbios Biosimulation Toolkit) and includes links to all sub-projects that make up SimTK.
Ontology: Application_Programming_Interface, Contact_Modeling, Linear_Algebra_Tool, Mechanical_Simulation, Molecular_Dynamics, Molecular_Interaction, Molecular_Model, Multibody_Dynamics, Numerical_Integrator, Numerical_Method, Optimizer
Keywords: computational biomechanics, gpu acceleration, linear algebra, molecular dynamics, multibody dynamics, numerical integration, numerical methods, optimization, random numbers, root finding

OpenMM

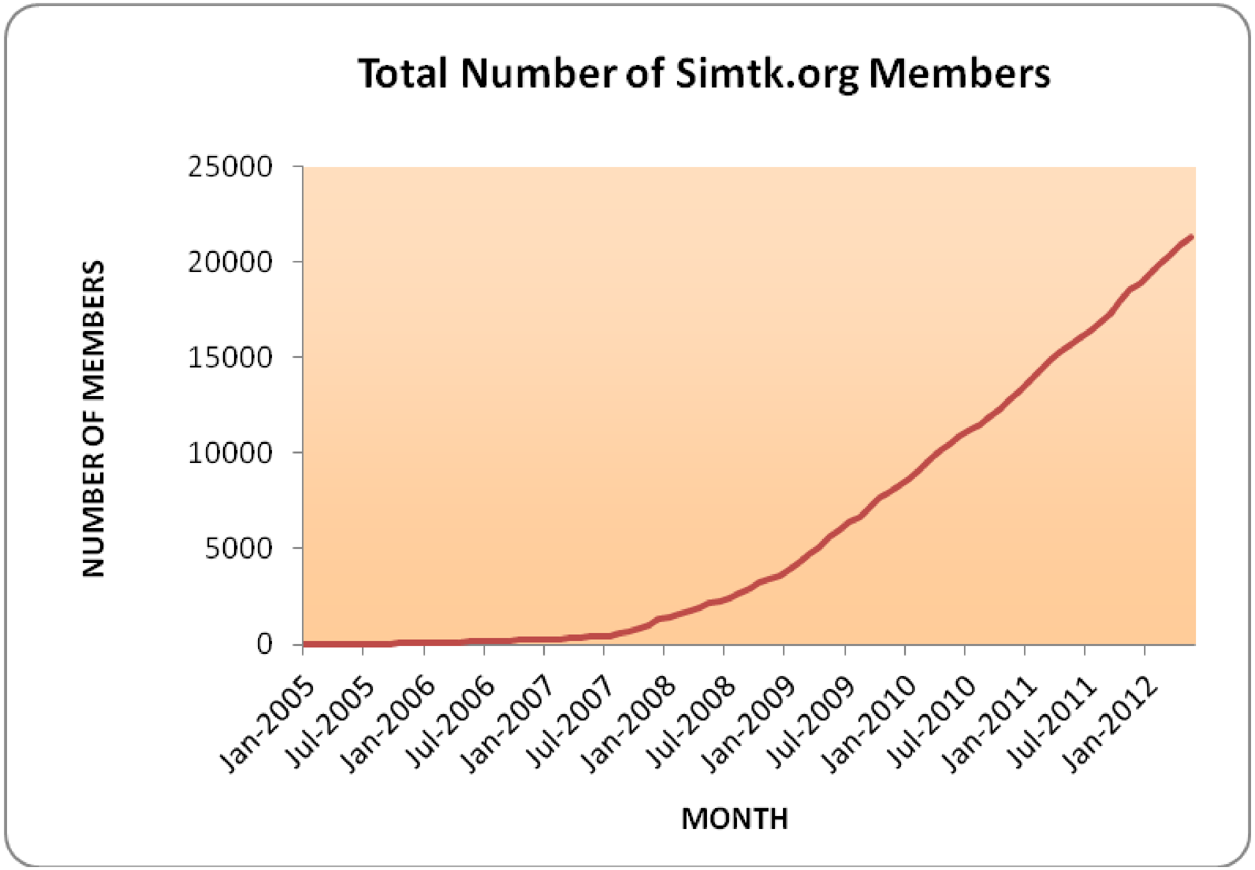
People

(affiliated with projects from your search results) Click on a person to see their profile and projects.

Adam Beberg
Adrien Coulet
Ahmet Bakan
Ahmet Erdemir
Ajay Seth
AJ Rader
Alain Laederach
Alain Laederach
Alan Grossfield
Alan Yen
Alberto Figueroa
Alena Shmygelska
Alex Limpaecher
Alex TenEyck
Allison Arnold
Allison Cohen
Allison Hall
Amanda Knutson
Anantharaman

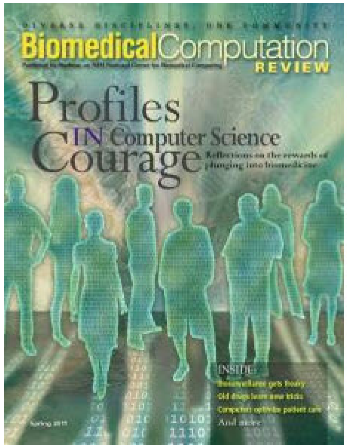
- Variety of projects hosted on Simtk.org
  - Software
  - Models & simulation data
  - Training materials
  - Publication-based projects
  - Projects for collaborations and more…

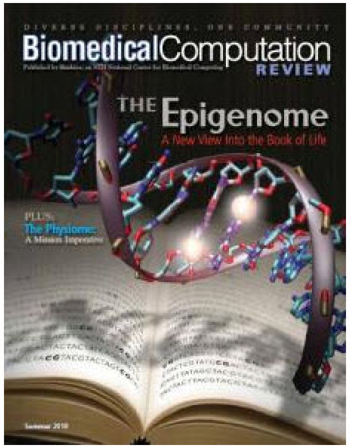
Simtk.org Statistics



- Over 21,900 members
- Average of 9200+ people visit each month
- Over 600 projects hosted on the site

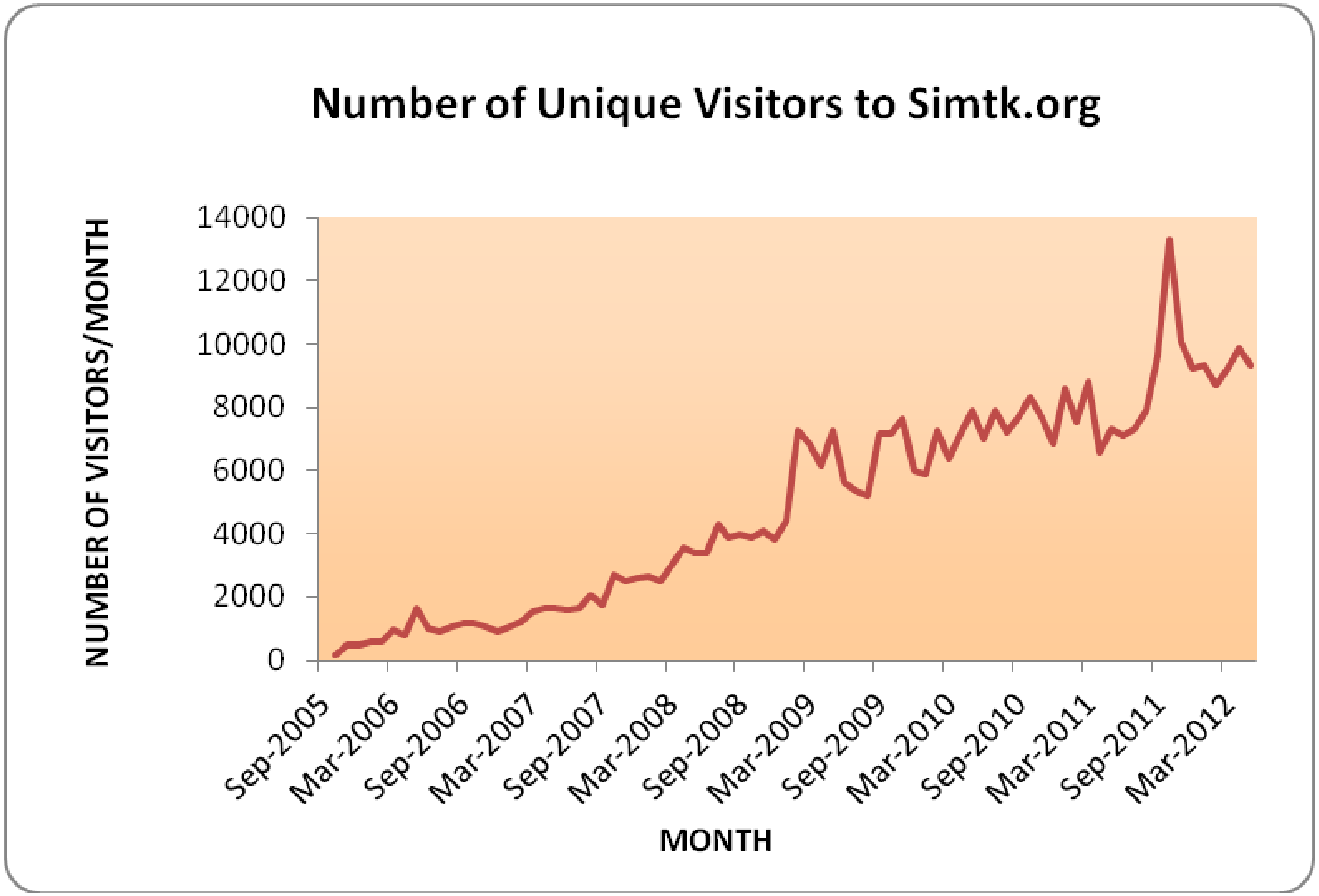
Biomedical Computation Review



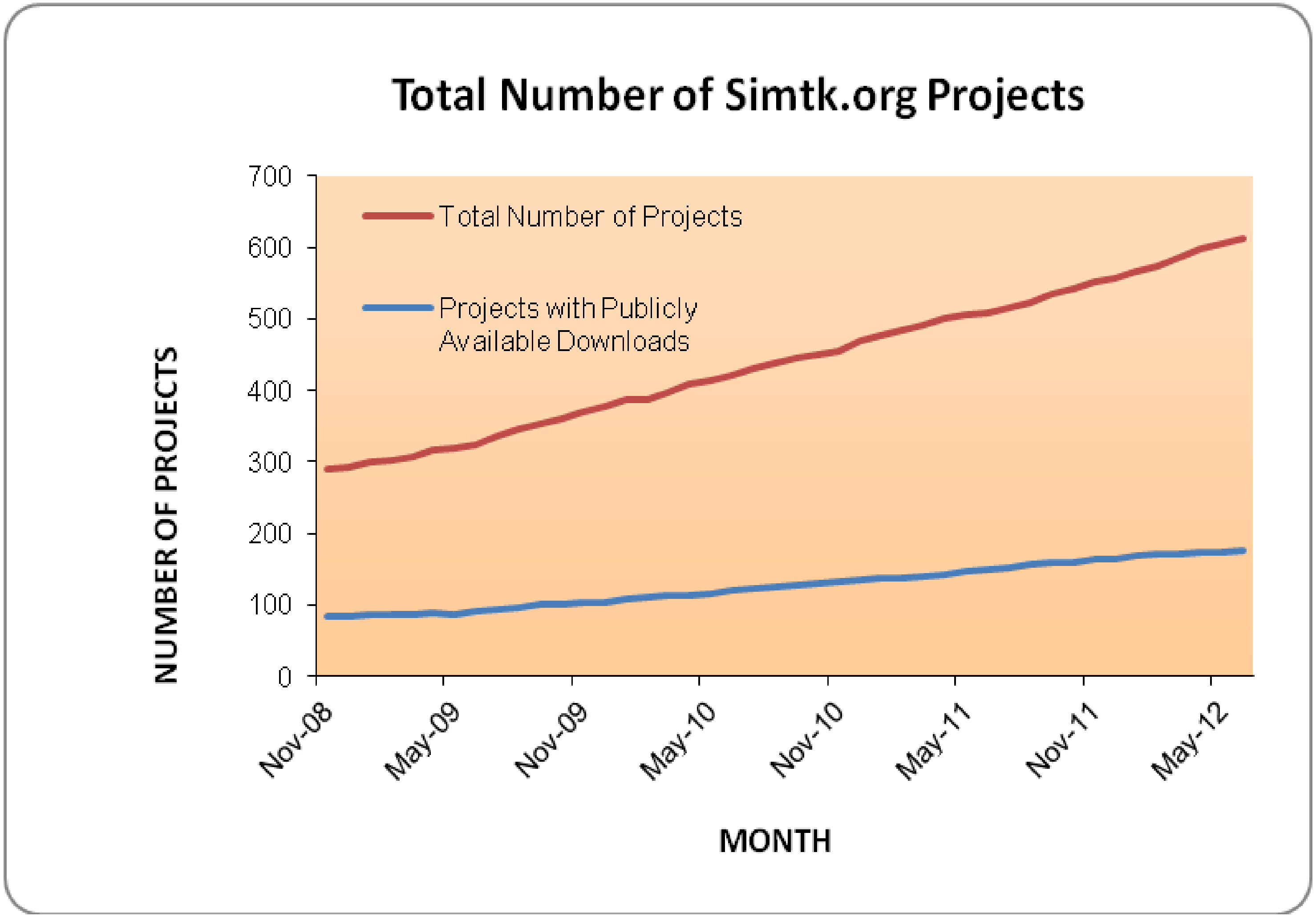


- ○ **Magazine on biomedical computation research, written for a general lay audience**
- ○ **Published 3 times a year**
- ○ **Also available on the web**
- ○ **Over 3600 subscribers to the print edition**

Simtk registered users



Simtk projects



Simtk project: core Simbios software

SimTKcore

Downloads

Some projects make source code available through the Subversion repository. You can check the Subversion Repository for source code for this project.

Examples programs using SimTK

Source code examples for rele4ase 1.5 of SimTKCore.

Release 1.5
August 8, 2008

Examples of programs that use SimTK core. Please download the SimTKExamples1.5.zip file after you install the correct version of SimTKcore for your platform. Please also make sure to read the notes provided in this folder: VisualStudioNotes.txt (windows), MakefileNotes.txt (Mac and linux) and ExampleAdenylateMobilitiesVMD_Notes.txt.
Release Notes

Download Links

Name	File Type	Platform	Updated
SimTKExamples1.5.zip (42 kB)	code	Any	Sep 05, 2008
Description: Includes 14 example programs and several important text files. Please see release description (above) for details.			
SimTKExamples1.5Updated.zip (75 kB)	code	Any	Sep 26, 2008
Description: Updated example files for Molmodel and musculoskeletal applications			

List all previous releases

Linux

SimTK core software for Linux

1.5.1
August 28, 2008

Patch to SimTK 1.5. Please download and read SimTKcore_README_Linux.pdf for installation instructions.
Release Notes

Project Administrator

Michael Sherman
Contact

Jack Middleton
Contact

Peter Eastman
Contact

Christopher Bruns
Contact

Blanca Pineda
Contact

Joy Ku
Contact

Team
10 members

Conclusions

- The three centers focusing on systems science and computational biology, substantial contributions
  - Scientific (papers, impact)
  - Community (people, meetings, magazines)
  - Infrastructure (software, data, models)

- The key success of the NCBC centers is producing and integrating all three of these in focused areas of science: in these cases = systems biology, biomedical informatics, physical simulation

Thanks.

russ.altman@stanford.edu

THEME B: Biomedical Image Post Processing under the NCBC Program

Ron Kikinis, M.D.

Robert Greenes Distinguished Director of Biomedical Informatics
Professor of Radiology, Harvard Medical School

Founding Director, Surgical Planning Laboratory, Brigham and Women’s Hospital
Principal Investigator, National Alliance for Medical Image Computing (a National Center for Biomedical Computing), and Neuroimage Analysis Center (a NIBIB National Resource Center)
Research Director, National Center for Image Guided Therapy

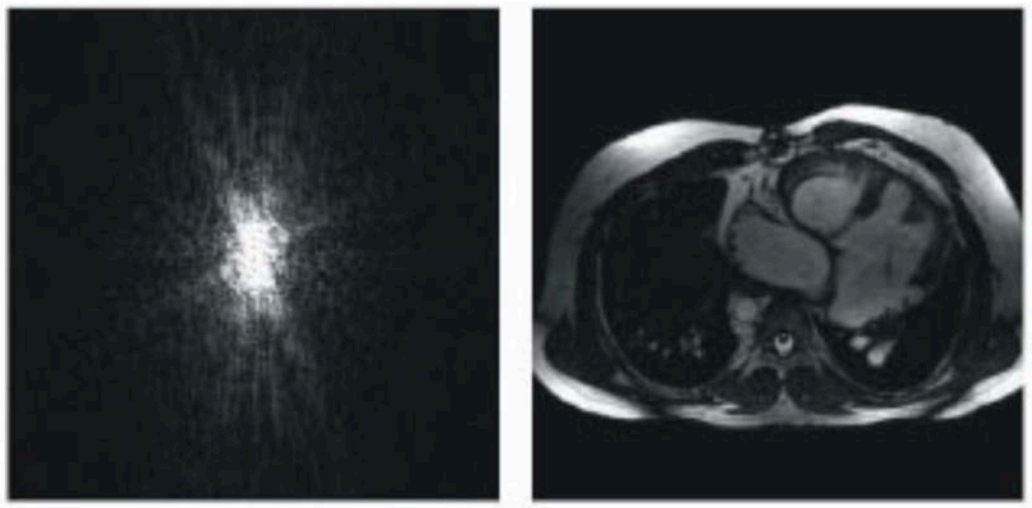
Charter

- A number of the NCBC centers have activities related to biomedical image analysis
  - The National Alliance for Medical Image Computing (NA-MIC) researches computational tools for the analysis and visualization of medical image data,
  - The Center for Computational Biology (CCB) is focused on the development of computational biological atlases of different populations, subjects, modalities, and spatio-temporal scales.
  - Several of the other centers have imaging related aims, including iDash, Simbios, and I2B2.

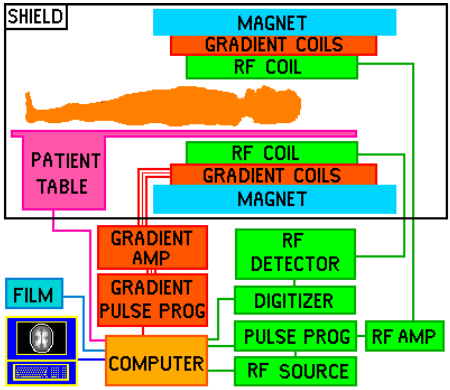
Introduction

- Imaging is complicated
- From physical measurement to image data.
  - MR: RF
  - CT: X-rays
  - Ultrasound
  - etc.

Example MRI
RF measurements
represented in k-space (left)
and
after Fourier transform
(right)



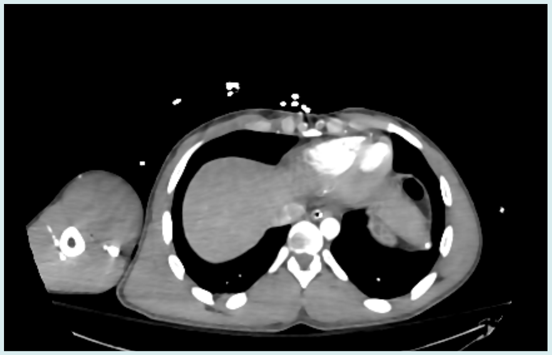
Source: http://www.biij.org/2008/1/e15/fig.asp?p=139&o=351



de.com/mysit

Post-Processing Is Needed

- From image data to information
- From information to knowledge



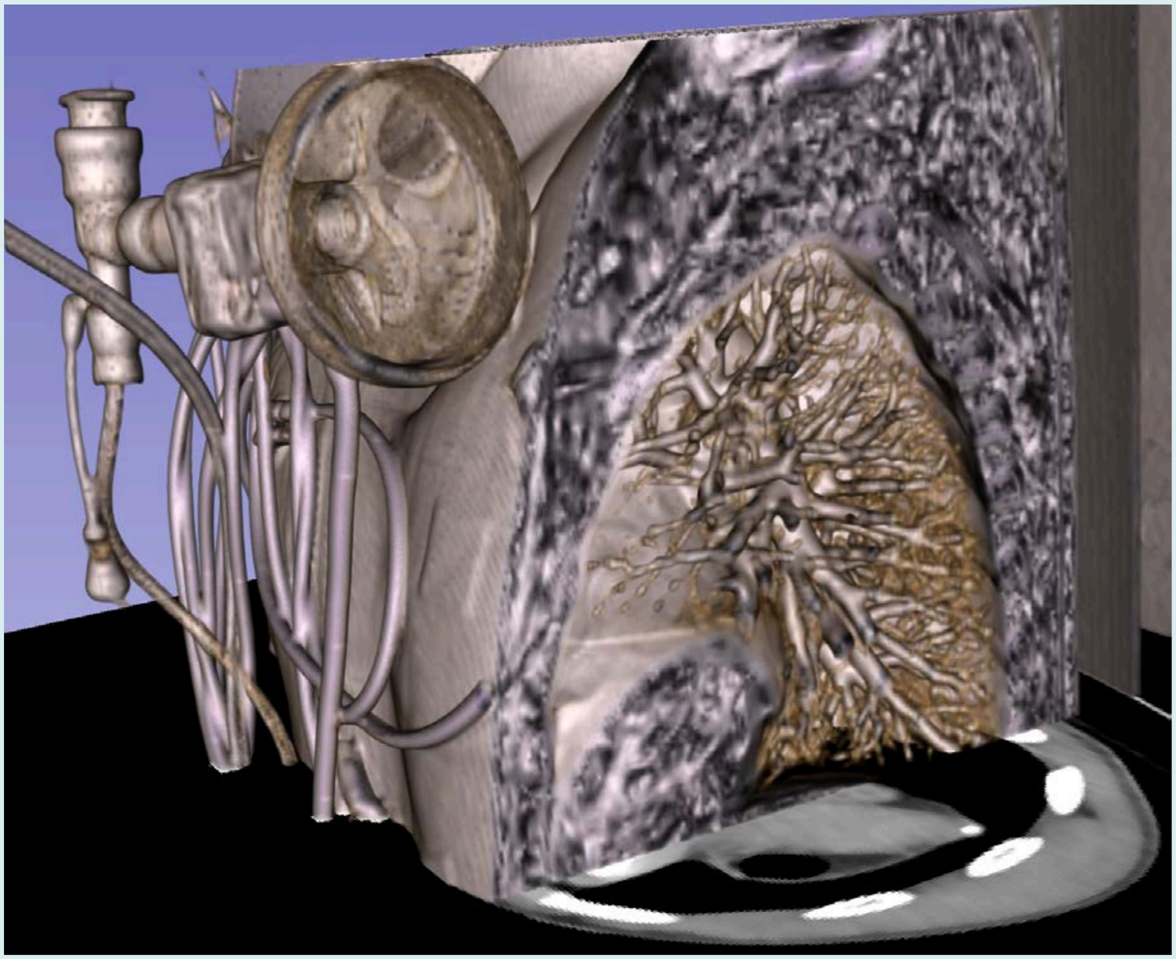
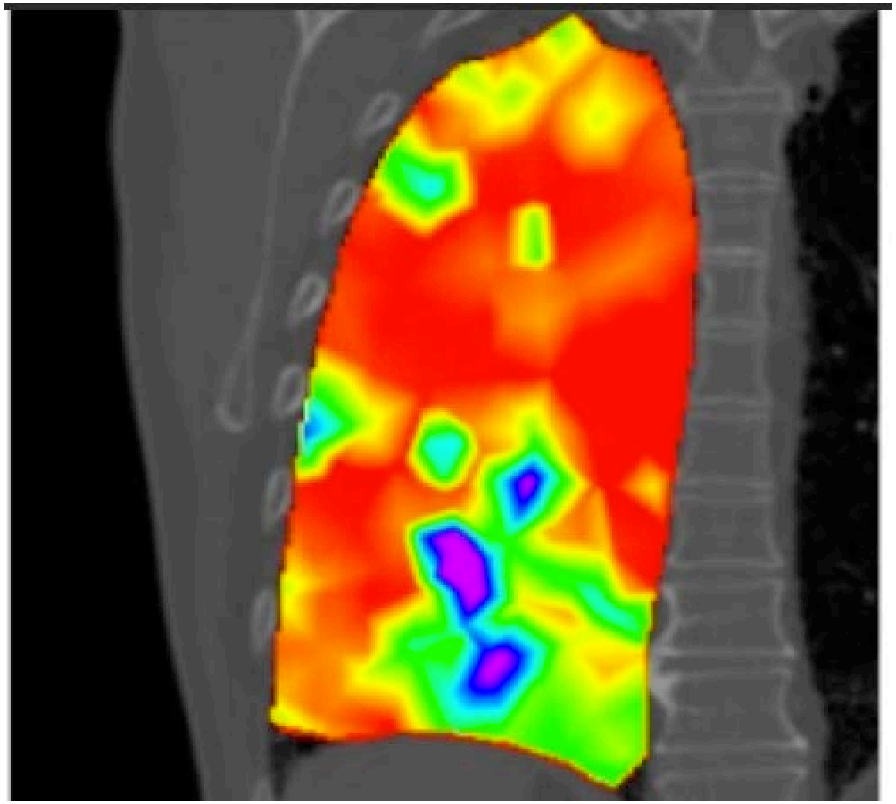


Image provided by Kikinis

Medical Image Computing (MIC)

- Imaging is of increasing importance in medicine, however:
  - More data
  - More complexity
    - fMRI, molecular imaging dMRI, 4DUS
  - More applications
    - Diagnosis
    - Therapy



Risholm P., Ross J., Washko G.R., Wells III W.M. *Probabilistic Elastography: Estimating Lung Elasticity*. Inf Process Med Imaging. 2011;22:699-710. PMID: 21761697. PMCID: PMC3249413

What Is Your Science?

- The basic science of one field is the applied science in a different field
- Image analysis concepts move
  - From applied math to computer vision to biomedical
- A caveat:
  - Clearly, your field of science is the most interesting to you.
  - Your colleagues think the same about their field of science

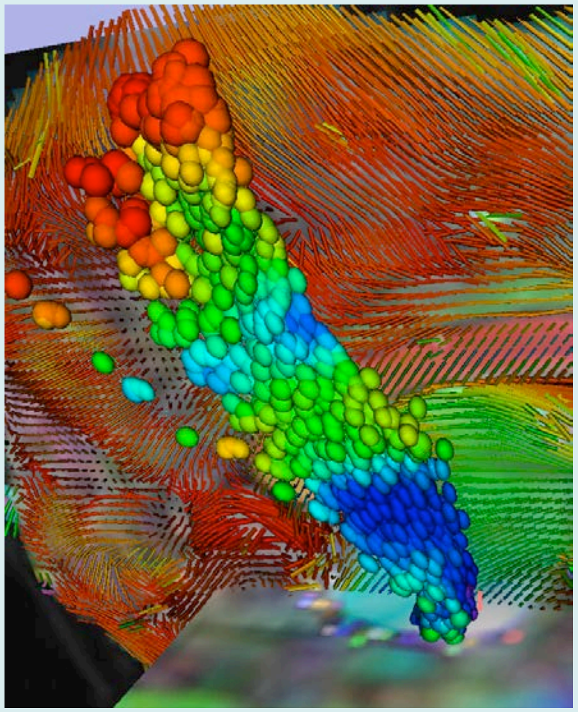


Image provided by Kikinis

From Algorithms To Tools

- Can it be done?
  - Prototypes
- **Is it worth doing?**
  - Tools for translational research
- Standard of care
  - Commercially available clinical “devices” with regulatory approval

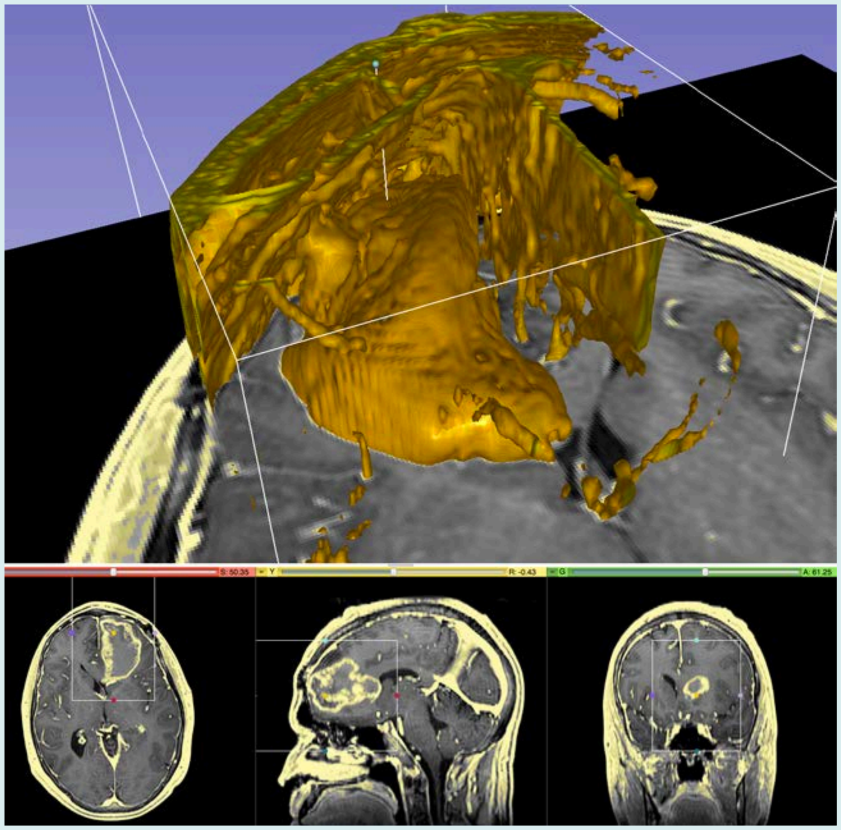


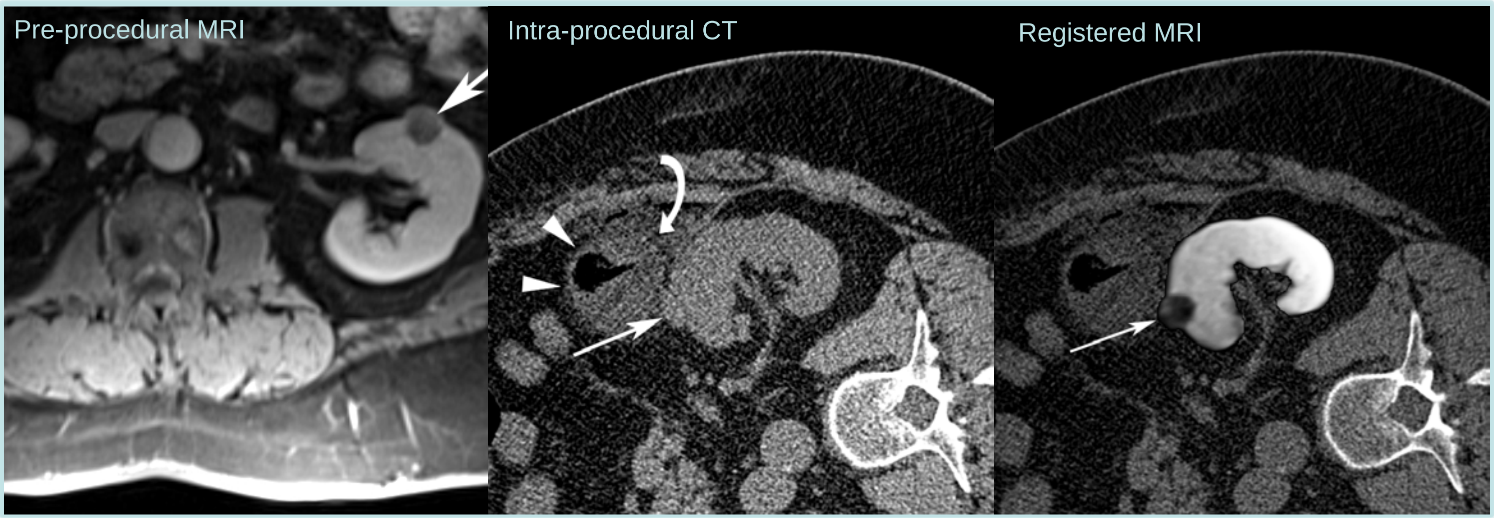
Image provided by Kikinis

Why Tools?

- A prototype will work until the completion of a thesis or paper
  - Not portable
  - Unstable
- A tool will work in your environment
  - Easy to install
  - Easy to use
  - Robust
- Significant software engineering investment is needed for the creation of robust tools

The Valley of Death

- Well-engineered software tools are needed for translational clinical research in image analysis
- Creating such tools is expensive and does not fit traditional funding mechanisms
  - Science agencies: its not innovative
  - Companies: no proven clinical utility, patents do not work, its too risky

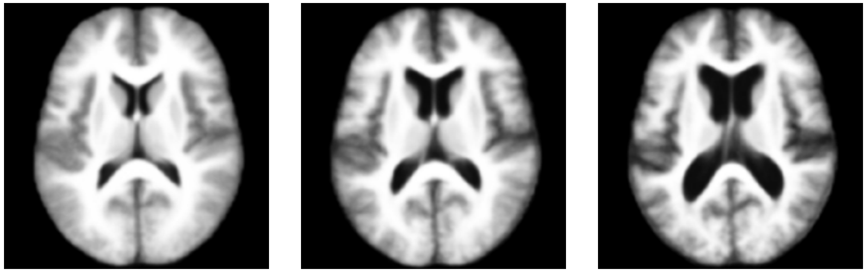


Oguro et al, Int J Comput Assist Radiol Surg, 2011
Elhawary et al, Acad Rad, 2010

Group Comparisons

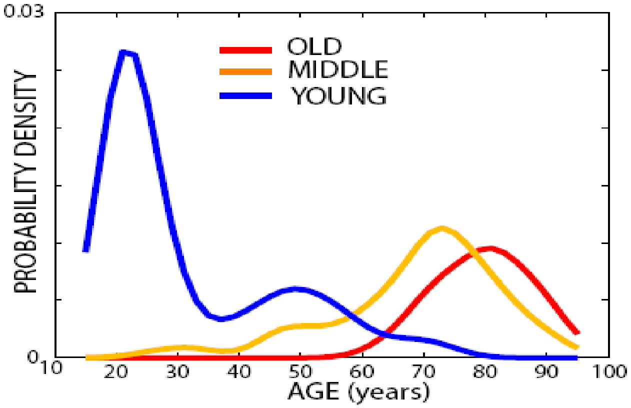
- Often used in basic imaging research
- Targets normal appearing structures
  - Typical appearance
  - Normal variability
- Extensive resources

M.R. Sabuncu, S.K. Balci, M.E. Shenton, and P. Golland. Image-Driven Population Analysis Through Mixture Modeling.
IEEE Transactions on Medical Imaging, 28(9):1473 - 1487, 2009



Young Middle Old

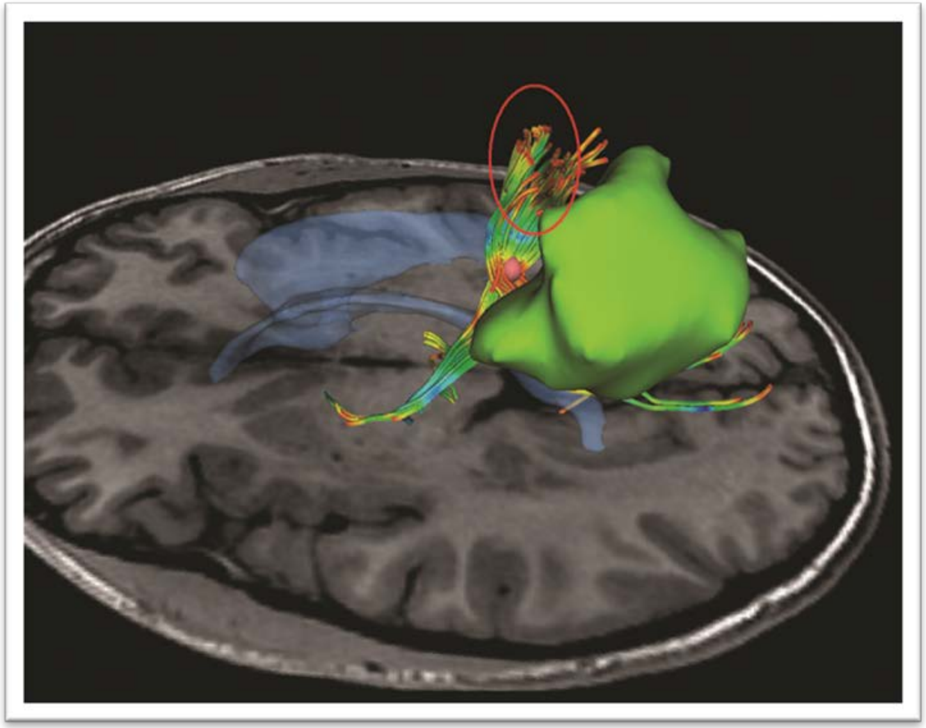
3 Templates



Golland, et al - MIT

Subject Specific Analysis

- Clinical motivation to understand the situation of one subject
- Targets pathology:
  - Where is the pathology?
  - What are important surrounding structures
- Limited resources



Golby A.J., Kindlmann G., Norton I., Yarmarkovich A., Pieper S., Kikinis R. *Interactive Diffusion Tensor Tractography Visualization for Neurosurgical Planning*. Neurosurgery. 2011 Feb; 68(2):496-505. PMID: 21135713

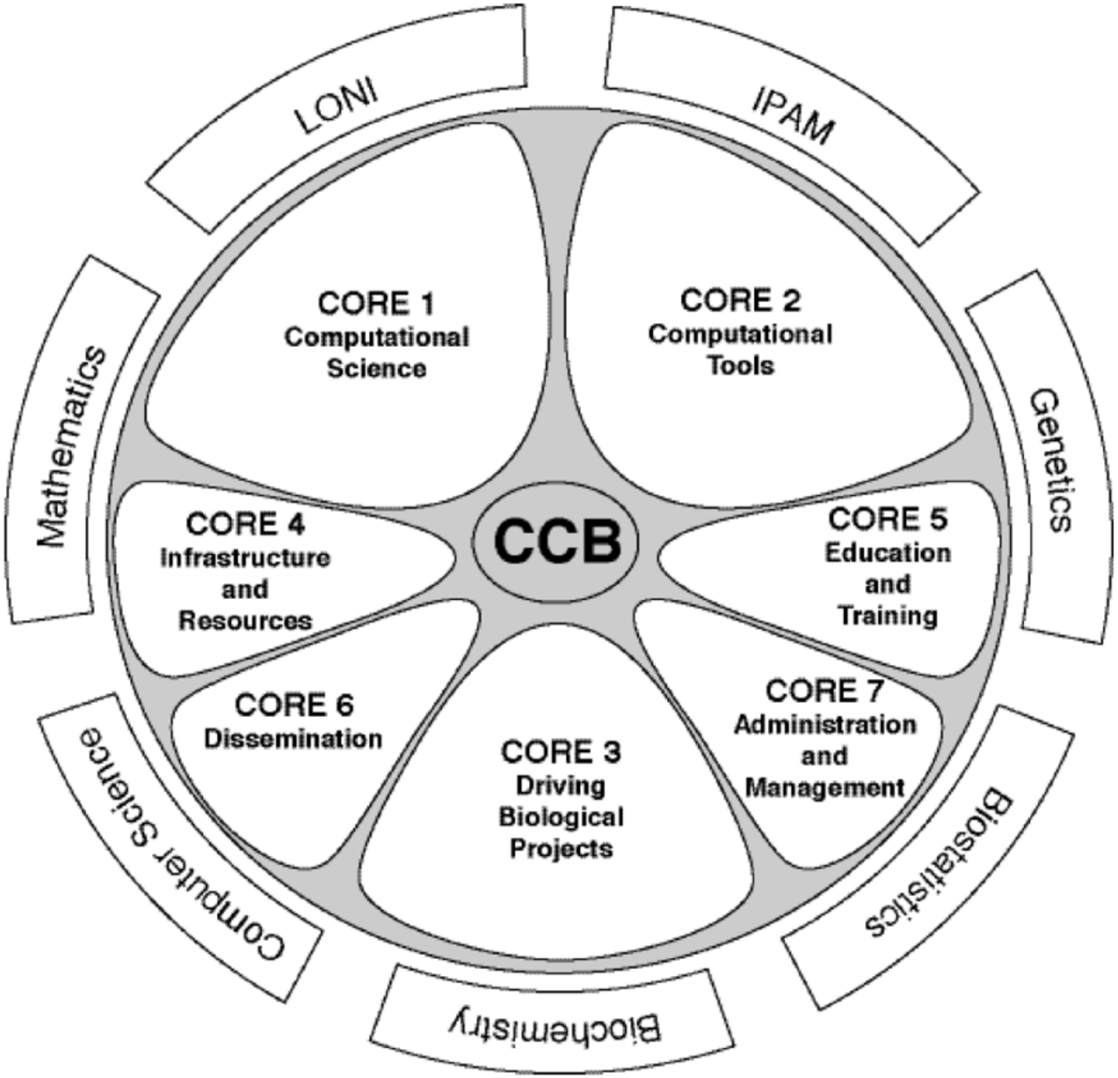
NCBC and Imaging

- CCB
- iDash
- Symbios
- I2B2
- NA-MIC

CCB

Center for Computational Biology, U54RR21813

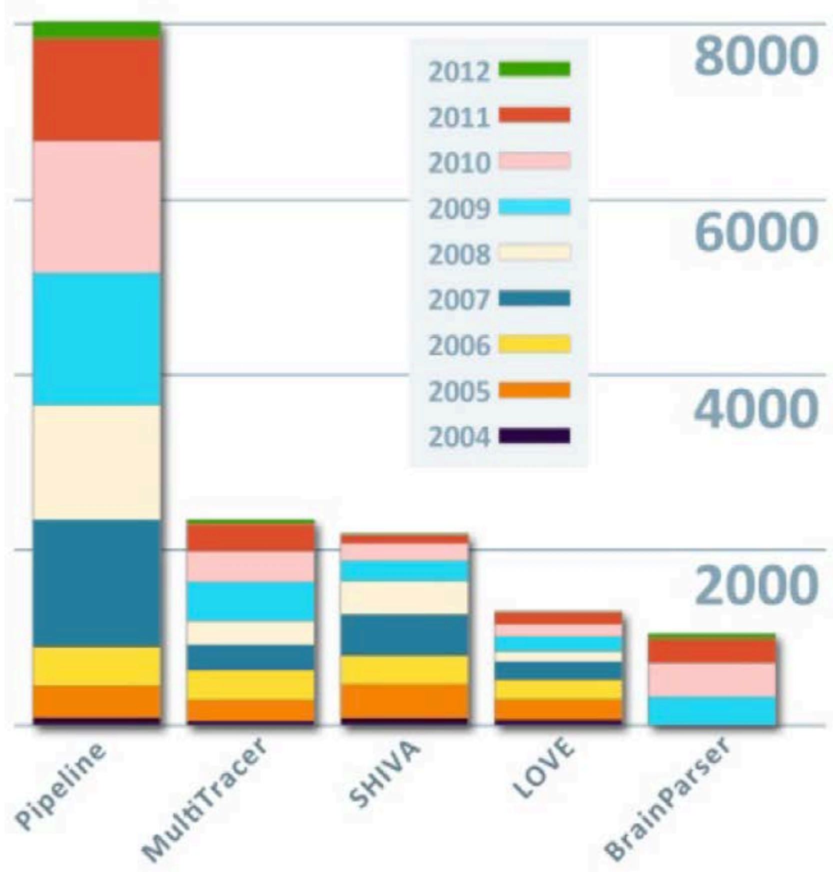
- Core 1: mathematical and computational research. Core 2: develops the tools to be applied to the four driving biological projects
- Core 3: DBPs: Mapping Genomic Function, Mapping Biological Structure and Mapping Brain Phenotype

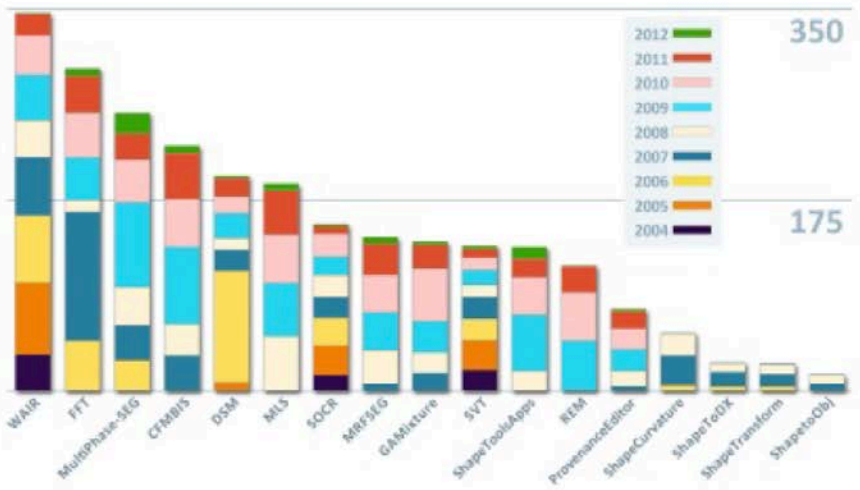


Content courtesy A. Toga

Accomplishments 20

- 280+ peer reviewed journal articles
- 70+ algorithmic prototypes and tools including:
  - Segmentation, registration, dMRI, workflow engines, provenance management, QA tools





Content courtesy A. Toga

Example tool: SVE

- The Segmentation Validation Engine (SVE) http://sve.loni.ucla.edu/, a segmentation challenge for skull-stripping.
- Registered users may download 40 MRI whole head volumes, process them and upload the results to the SVE server.
- Uploaded segmentation results are processed to produce a set of agreement measures based on comparison with manually delineated brain volumes.

Content courtesy A. Toga

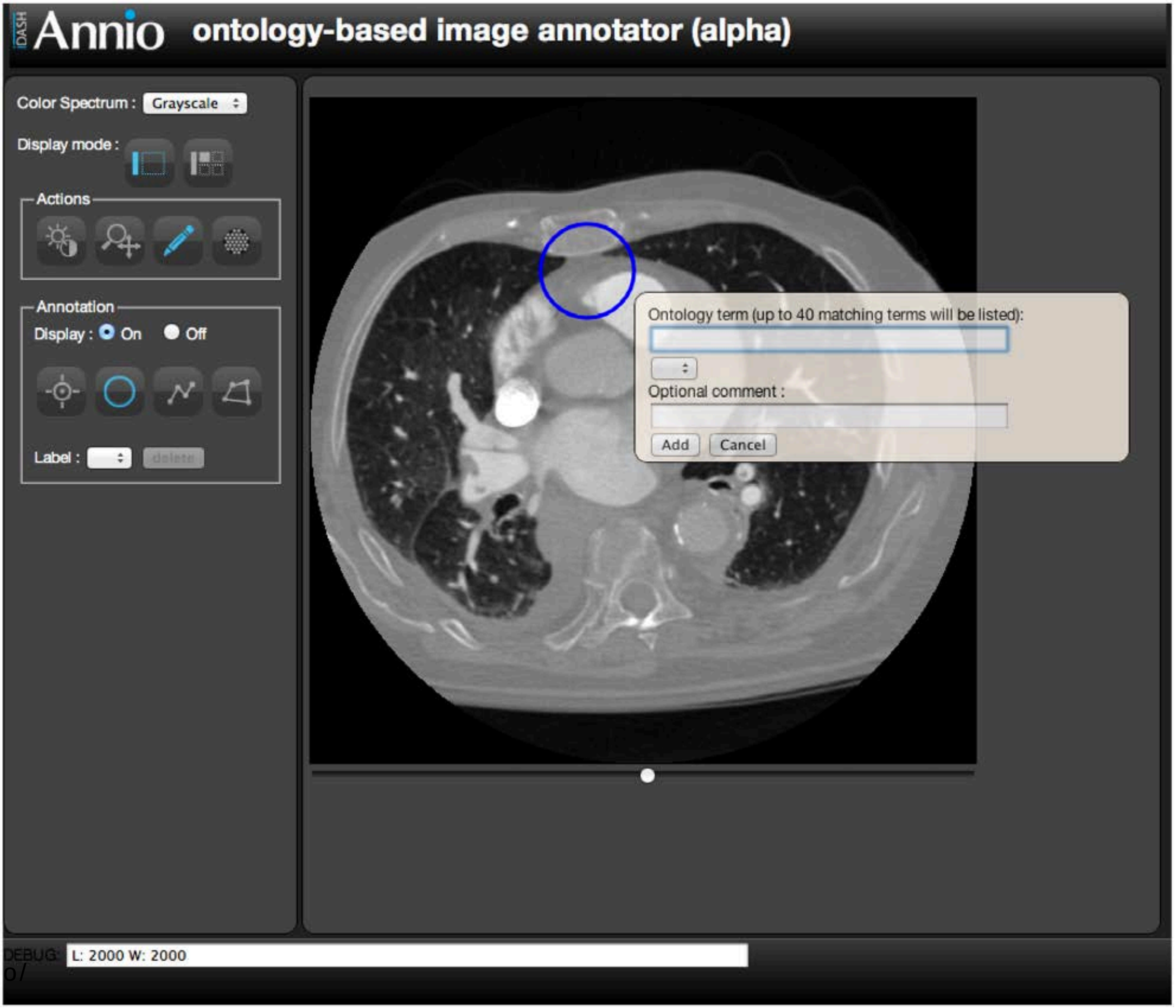
iDash

Integrating Data for Analysis, Anonymization and SHaring (iDASH)

- The iDASH center provides a secure, privacy-preserving environment in which researchers can analyze genomic, transcriptomic, clinical, behavioral, and social data relevant to health.

iDash

As part of the comprehensive support for annotation of clinical data, iDash is developing a web based image annotation tool.



Color Spectrum : Grayscale

Display : ● On ○ Off

Ontology term (up to 40 matching terms will be listed):

Optional comment :

DEBUG: L: 2000 W: 2000

http://imaging-idash.ucsd.edu/annio/

Simbios

Simbios is the NIH Center for physics-based Simulation of Biological Structures.
U54 GM072970,

Simulation-Based Treatment Planning

Experimental Data

Biomechanical Models

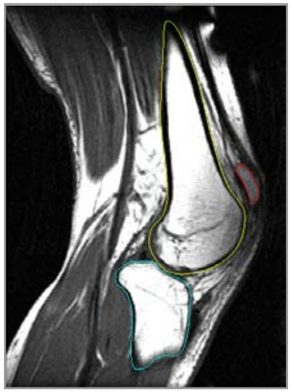
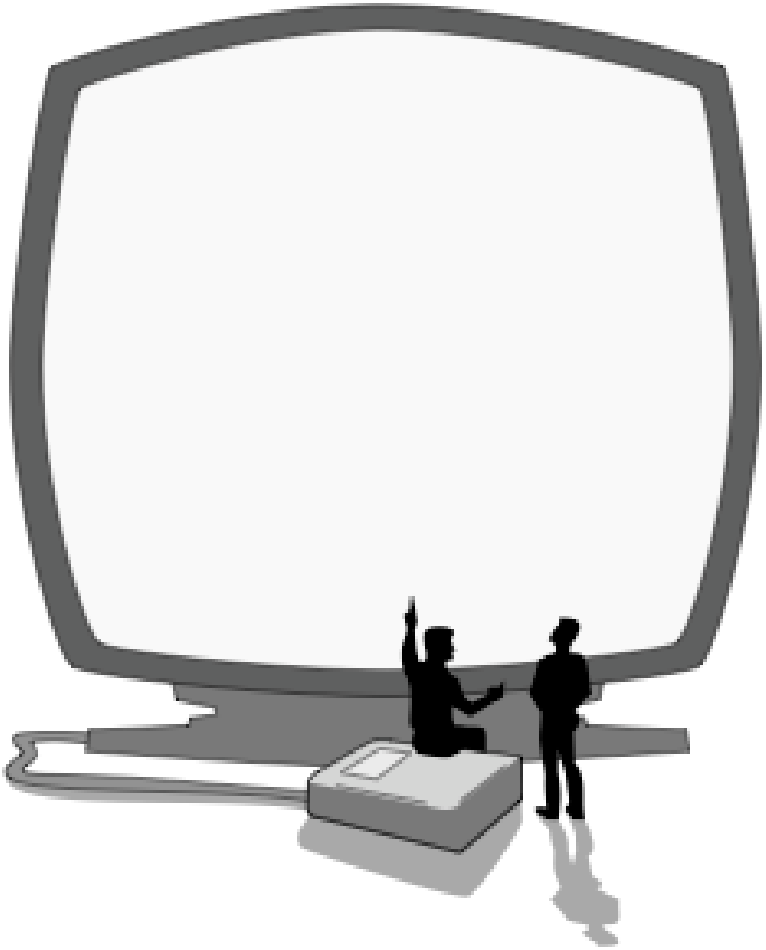


Image analysis



I2B2

Informatics for Integrating Biology and the Bedside

- The i2b2 Center is developing a scalable informatics framework that will enable clinical researchers to use existing clinical data for discovery research
- Massive data mining in clinical databases

mi2b2

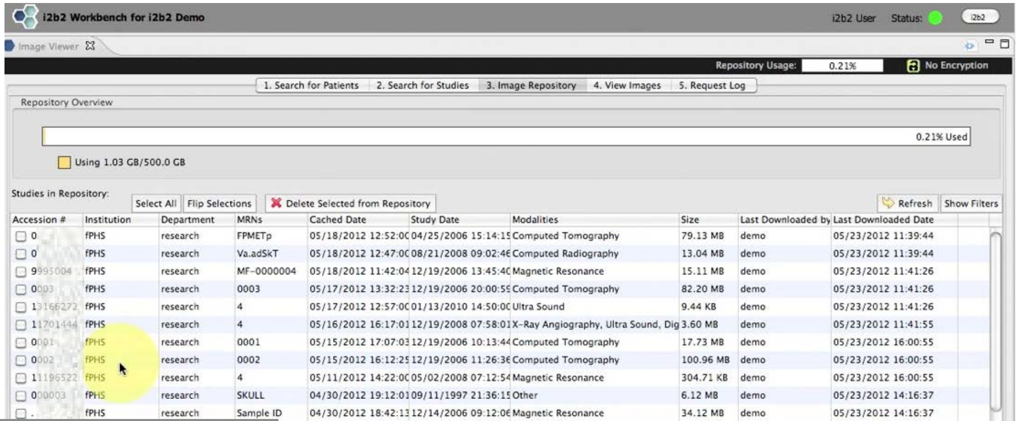
Medical Image Informatics Bench to Bedside

Massive Image Data Mining in Clinical PACS systems

- Integrates Informatics Databases (RPDR/i2b2) with PACS
- DICOM Query/Retrieve to Desktop
- Logged Access and IRB Controls
- Governed Access to PACS
  - Radiology Department Controls Access Time and Rate
  - Prevent Interference with Clinical Work
- Free and Open Source http://www.mi2b2.org

mi2b2

IRB-Approved Medical Records Queries Identify Studies of Interest

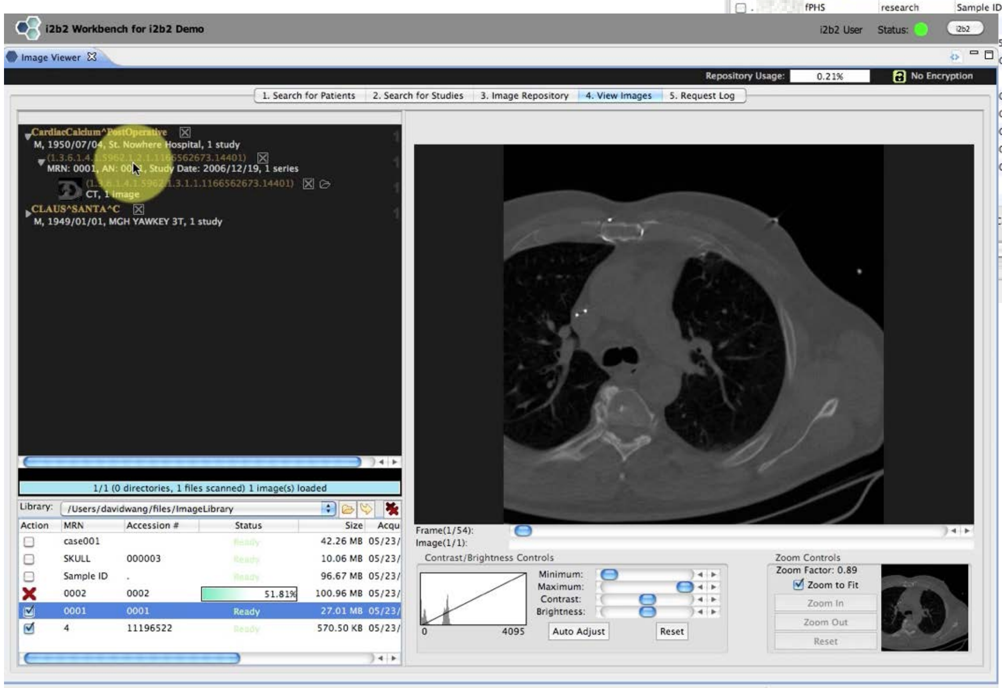


Using 1.03 GB/500.0 GB

Studies in Repository:

Accession #	Institution	Department	MRNs	Cached Date	Study Date	Modalities	Size	Last Downloaded by	Last Downloaded Date
☐ 0[redacted]	fPHS	research	FPMETp	05/18/2012 12:52:0	04/25/2006 15:14:1	Computed Tomography	79.13 MB	demo	05/23/2012 11:39:44
☐ 0[redacted]	fPHS	research	Va.adSkT	05/18/2012 12:47:0	08/21/2008 09:02:4	Computed Radiography	13.04 MB	demo	05/23/2012 11:39:44
☐ 9[redacted]	fPHS	research	MF-0000004	05/18/2012 11:42:04	12/19/2006 13:45:4	Magnetic Resonance	15.11 MB	demo	05/23/2012 11:41:26
☐ 0[redacted]	fPHS	research	0003	05/17/2012 13:32:2	12/19/2006 20:00:5	Computed Tomography	82.20 MB	demo	05/23/2012 11:41:26
☐ 1[redacted]	fPHS	research	4	05/17/2012 12:57:0	01/13/2010 14:50:0	Ultra Sound	9.44 KB	demo	05/23/2012 11:41:26
☐ 1[redacted]	fPHS	research	4	05/16/2012 16:17:01	12/19/2008 07:58:01	X-Ray Angiography, Ultra Sound, Dig	3.60 MB	demo	05/23/2012 11:41:55
☐ 0[redacted]	fPHS	research	0001	05/15/2012 17:07:03	12/19/2006 10:13:44	Computed Tomography	17.73 MB	demo	05/23/2012 16:00:55
☐ 0[redacted]	fPHS	research	0002	05/15/2012 16:12:2	12/19/2006 11:26:3	Computed Tomography	100.96 MB	demo	05/23/2012 16:00:55
☐ 1[redacted]	fPHS	research	4	05/11/2012 14:22:0	05/02/2008 07:12:54	Magnetic Resonance	304.71 KB	demo	05/23/2012 16:00:55
☐ 0[redacted]	fPHS	research	SKULL	04/30/2012 19:12:01	09/11/1997 21:36:1	Other	6.12 MB	demo	05/23/2012 14:16:37
☐ .[redacted]	fPHS	research	Sample ID	04/30/2012 18:42:1	12/14/2006 09:12:0	Magnetic Resonance	34.12 MB	demo	05/23/2012 14:16:37
				04/30/2012 18:37:0	12/19/2006 10:13:44	Magnetic Resonance	19.79 MB	demo	05/23/2012 14:16:37
			5	04/30/2012 18:32:03	10/14/2005 12:52:5	Magnetic Resonance	13.77 MB	demo	05/18/2012 12:17:55
			032	04/30/2012 18:27:0	12/19/2006 12:25:2	Magnetic Resonance	31.96 KB	demo	05/11/2012 13:39:10
				04/30/2012 18:22:17	12/19/2006 13:53:32	Computed Tomography	76.92 MB	demo	05/11/2012 13:39:12
			031	04/30/2012 18:17:0	12/19/2006 12:25:2	Magnetic Resonance	32.01 KB	demo	05/11/2012 13:39:10
			028	04/30/2012 18:07:0	12/19/2006 12:25:2	Magnetic Resonance	32.20 KB	demo	05/11/2012 14:21:48
			010	04/30/2012 18:02:02	12/19/2006 00:00:1	Magnetic Resonance	7.57 MB	demo	05/18/2012 12:17:55
			013	04/30/2012 17:57:0	12/19/2006 00:00:1	Magnetic Resonance	855.69 KB	demo	05/18/2012 12:32:28
			008	04/30/2012 17:37:03	12/19/2006 00:00:1	Magnetic Resonance	10.18 MB		
				04/30/2012 17:32:0	12/19/2006 11:11:54	Computed Tomography	219.93 KB		
				04/30/2012 17:27:0	12/19/2006 11:25:24	Computed Tomography	2.51 MB	demo	05/08/2012 10:30:56

Cache to XNAT (Please select one or more studies.)



CardiacCalcium^PostOperative
M, 1950/07/04, St. Nowhere Hospital, 1 study
(1.3.6.1.4.1.5962.1.2.1.1166562673.14401)
MRN: 0001, AN: 0001, Study Date: 2006/12/19, 1 series
(1.3.6.1.4.1.5962.1.3.1.1.1166562673.14401)
CT, 1 image
CLAUS^SANTA^C
M, 1949/01/01, MGH YAWKEY 3T, 1 study

1/1 (0 directories, 1 files scanned) 1 image(s) loaded

Library: /Users/davidwang/files/ImageLibrary

Action	MRN	Accession #	Status	Size	Acqu
☐	case001		Ready	42.26 MB	05/23/
☐	SKULL	000003	Ready	10.06 MB	05/23/
☐	Sample ID	.	Ready	96.67 MB	05/23/
☒	0002	0002	51.81%	100.96 MB	05/23/
☑	0001	0001	Ready	27.01 MB	05/23/
☑	4	11196522	Ready	570.50 KB	05/23/

Frame(1/54):
Image(1/1):

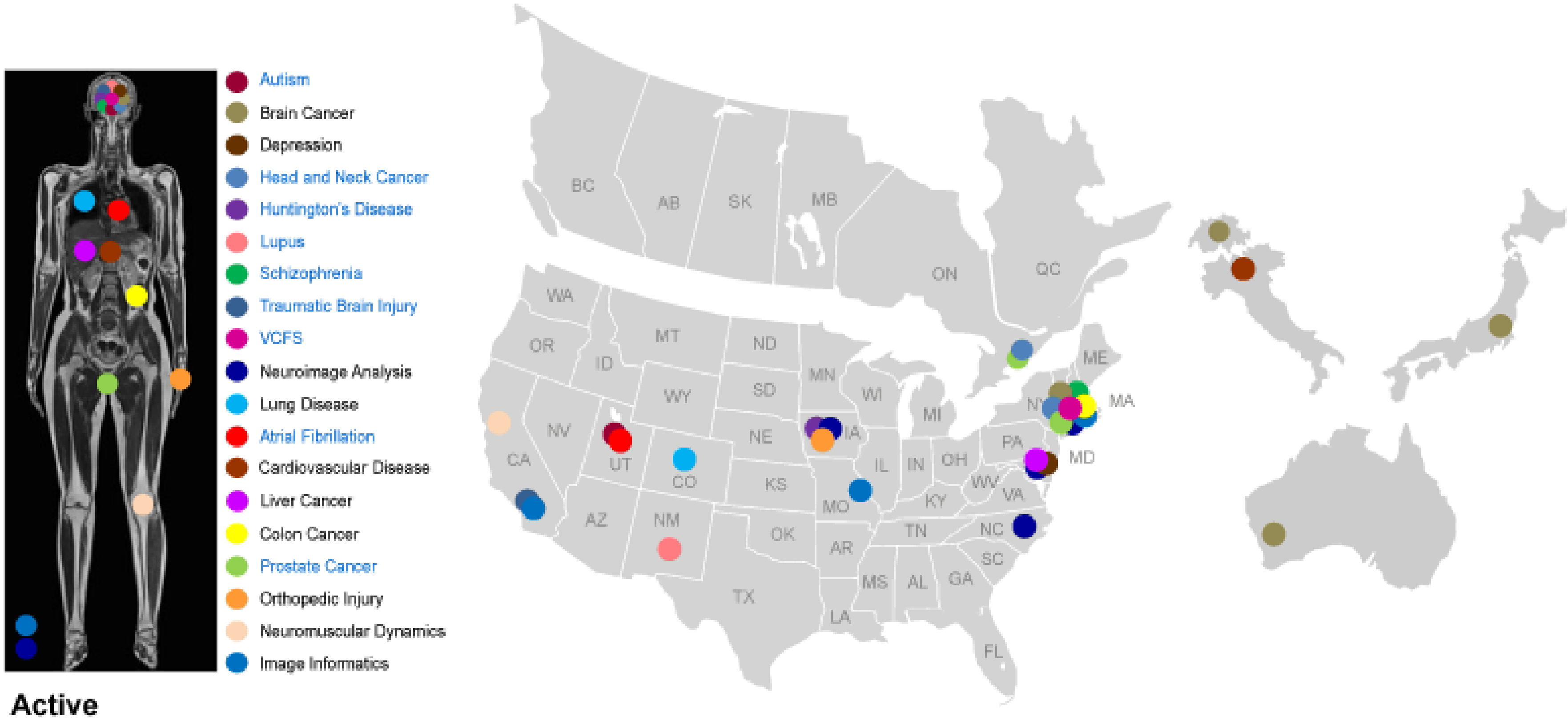
Zoom Factor: 0.89
☑ Zoom to Fit

Selected Studies are Downloaded to Local Storage for Further Research Processing

NA-MIC

- The National Alliance for Medical Image Computing (NA-MIC), is a community of esearchers
- Focus on
  - Subject specific image analysis
  - NA-MIC kit, including 3D Slicer as a platform for dissemination

NA-MIC Community



- ● Autism
- ● Brain Cancer
- ● Depression
- ● Head and Neck Cancer
- ● Huntington's Disease
- ● Lupus
- ● Schizophrenia
- ● Traumatic Brain Injury
- ● VCFS
- ● Neuroimage Analysis
- ● Lung Disease
- ● Atrial Fibrillation
- ● Cardiovascular Disease
- ● Liver Cancer
- ● Colon Cancer
- ● Prostate Cancer
- ● Orthopedic Injury
- ● Neuromuscular Dynamics
- ● Image Informatics

Active

● R01MH084795
● R01EB005973
● UL1RR025758
● Mario Negri Institute, Italy
●● U41RR019703
● U54EB005149-05S2
● U54LM008748
● NSF CCF-0916526
● U54GM072970
● U24RR025736
● CO-ME, Switzerland
● R01EB008171
● P41RR013218
● U24RR021992
● U01HL089897
● R01EB006733
● U24RR021382
●● OCAIRO, Canada
● R01CA124377
● R01NS050568
● U24RR026057
● R01CA131718
● R21EB009900
● AIST, Japan
● R01CA11128
● U54EB005149-05S3
● UWA, Australia

Completed

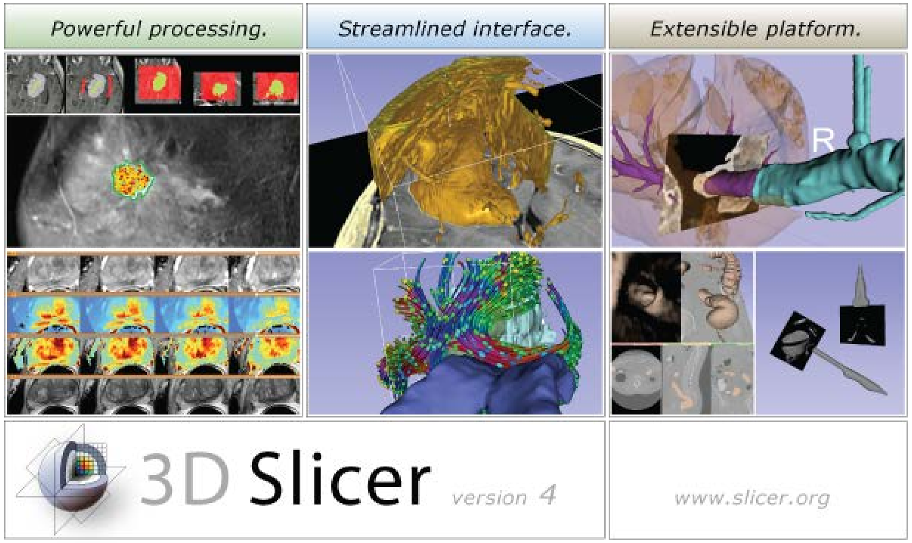
● U54EB005149-04S1

Picture 2011, courtesy Kapur, Jakab, Kikinis

Why Open Source?

- Collaborate freely
  - Proprietary software allows collaboration only among customers of that package
  - Neutral territory in multi-vendor scenarios,
- Reduce duplication
- Extensible

3D Slicer

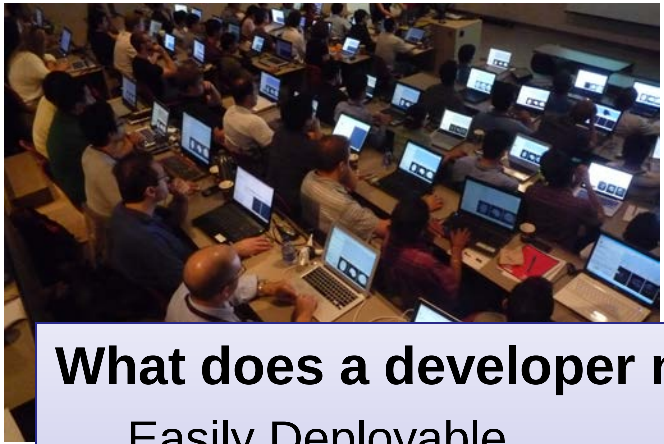


- 3D Slicer bridges the “valley of death” for subject specific analysis
  - Free open source software
    - Enables scientific collaboration
    - License allows painless translation to proprietary clinical tools
  - Well-engineered high-performance core
    - Software engineering methodology, multi-platform
  - Many options for extensions

3DSlicer

Picture courtesy Wendy Plesniak

Slicer for Translational Research



3DSlicer

What does a developer need ?

- Easily Deployable
- Extensible and Reconfigurable
- Rich Utility Libraries
- Stable Base

What does a user expect ?

- Easy Install and Upgrade
- "Standard" Clinical Behavior
- Advanced Functionality
- Consistent Interface

3D Slicer: a cross platform system for ***translating*** innovative algorithms into clinical research applications

Courtesy R. Kikinis

Slicer Is Open

- Open Science
  =
  Open Source
  +
  Open Data
  +
  Open Community

Tokyo 2010, hosted by H. Iseki

Boston 2010

Slicer Is Accessible

- Designed from scratch for:
  - Multi-platform support
  - Multi-site development
- International developer community
- International user community

Slicer is Extensible

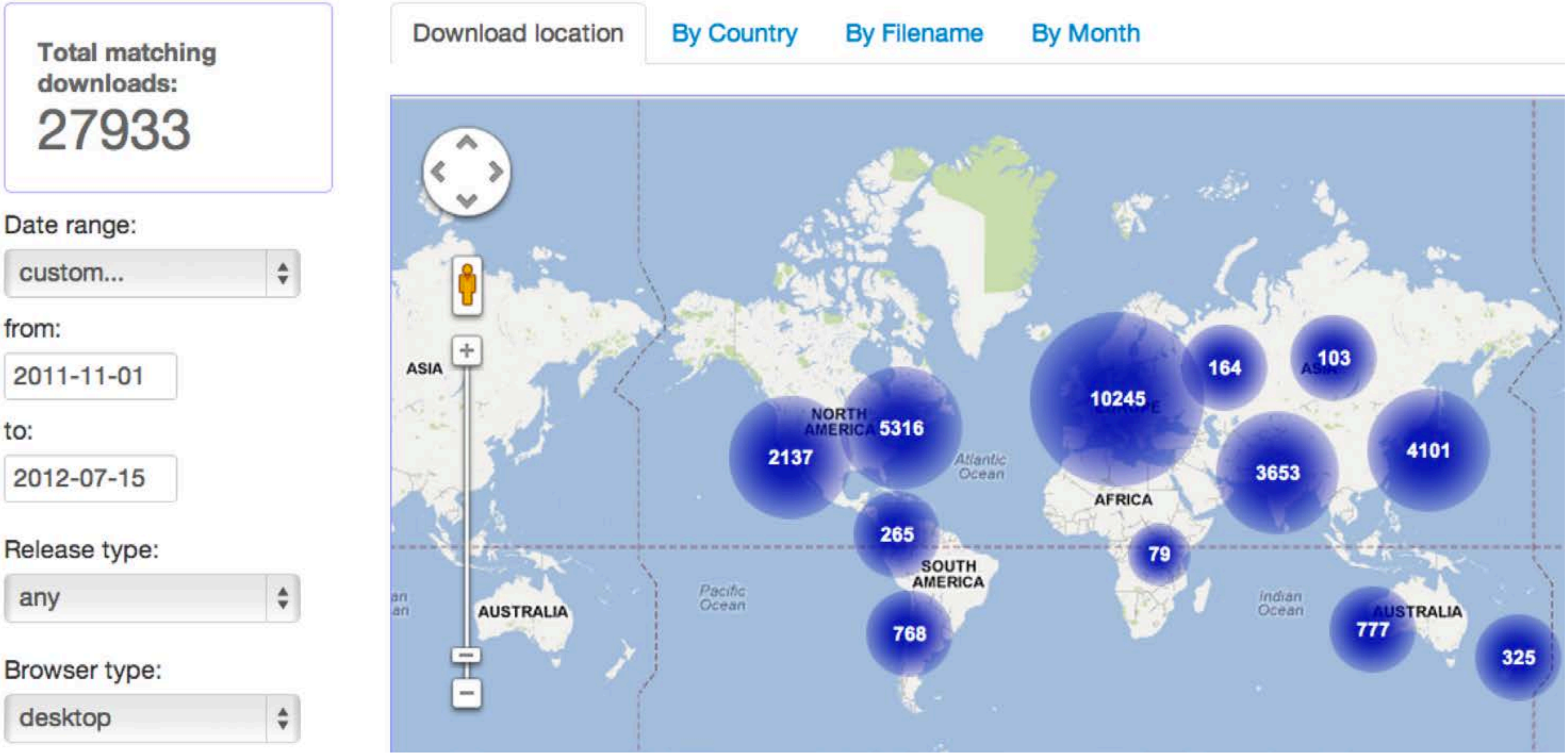
- Variety of options for the developers:
  - Developer startup time and control of Slicer

    Less ↓ More

    - **Scripted module**: TCL or Python scripts
      - simple, no compilation needed
      - limited access to Slicer internals
    - **Command-line module**: .exe file
      - simple, executable without Slicer
      - no access to Slicer internals, Slicer compilation needed
    - **Loadable (interactive) module**: .dll
      - full access to Slicer internals
      - Slicer compilation needed, requires Slicer core knowledge

Slicer 4 Downloads

http://download.slicer.org



Total matching downloads: 27933

Date range: custom...

from: 2011-11-01

to: 2012-07-15

Release type: any

Browser type: desktop

Picture courtesy M. Halle, M. Jakab, T. Kapur

URL's

National Alliance for Medical Image Computing
www.na-mic.org

Neuroimage Analysis Center
Nac.spl.harvard.edu

Surgical Planning Laboratory,
Brigham and Women's Hospital
spl.harvard.edu

National Center For Image Guided Therapy
www.ncigt.org

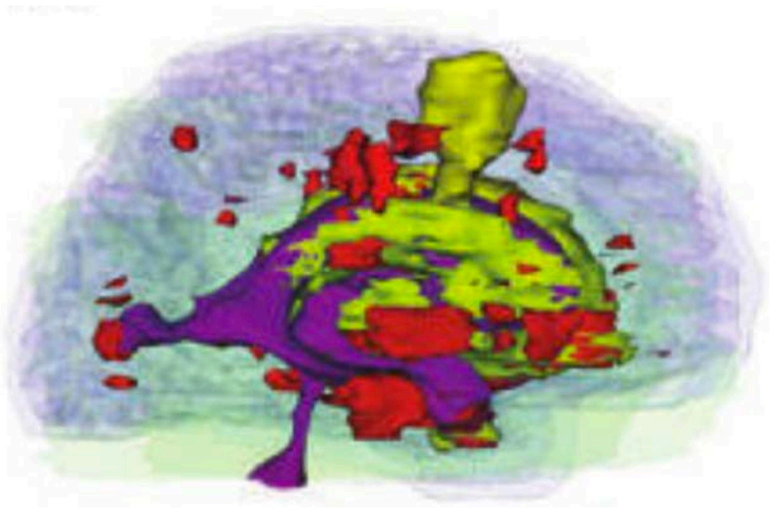
Subject Specific Analysis

- Studies focused on subject specific analysis are an example of the translational science paradigm discussed previously. The goal is to develop capabilities to analyze:
  - One subject
  - Short time for the analysis
- Limited automation
- Limited computational resources

- Example:
  - TBI

NA-MIC DBP on TBI

Team:
BWH, Utah, Kitware, UNC,
Boston Univ., UCSD



Irimia A., Chambers M.C., Alger J.R., Filippou M., Prastawa M.W., Wang B., Hovda D., Gerig G., Toga A.W., Kikinis R., Vespa P.M., Van Horn J.D. *Comparison of Acute and Chronic Traumatic Brain Injury using Semi-automatic Multimodal Segmentation of MR Volumes*. J Neurotrauma. 2011 Nov;28(11):2287-306

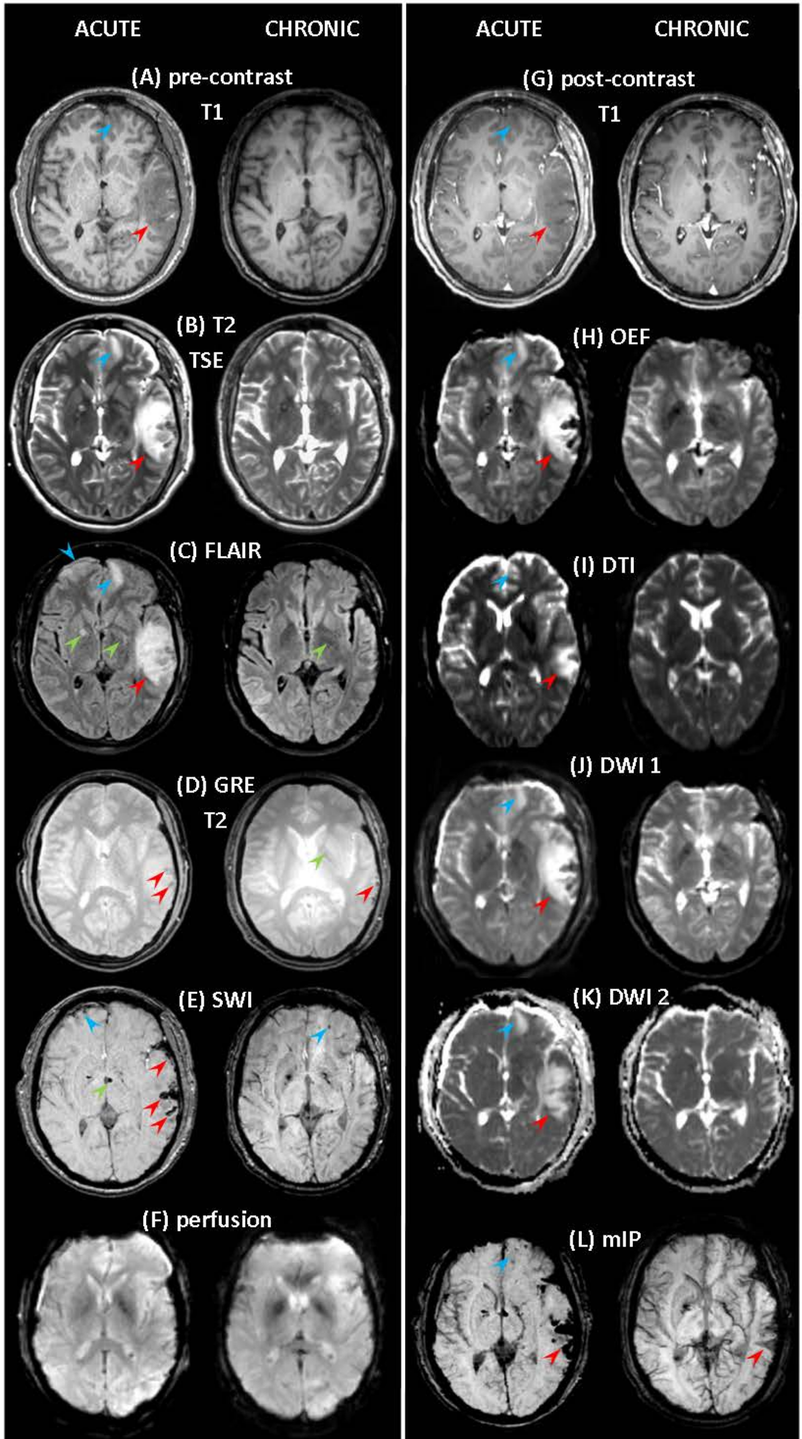
- Biomedical image analysis
- Segmentation
- Processing
- Modeling
- Measurement
- Visualization

http://www.na-mic.org/

Example Case

Longitudinal neuroimaging of TBI using multimodal MRI at the Brain Injury Research Center (BIRC), UCLA

- Acute imaging within 1st 48 hours after admission
- Chronic imaging 6 months post injury
- Multiple sequences

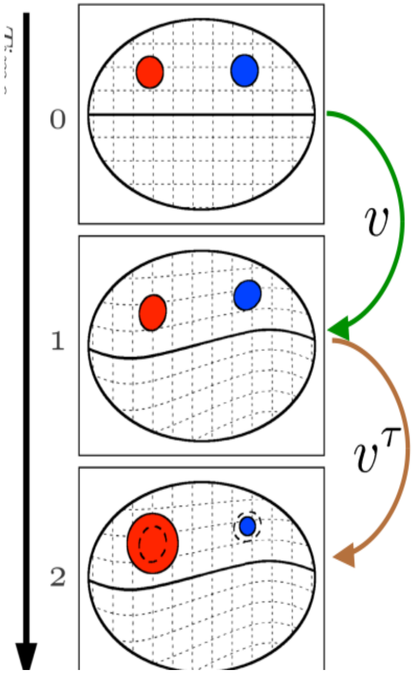


P. Vespa, J. Alger, et al.

Algorithm Development 1

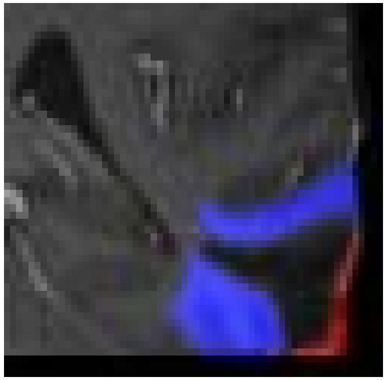
Registration:
Geometric 'Metamorphosis'

Joint estimation of global background & pathology deformation



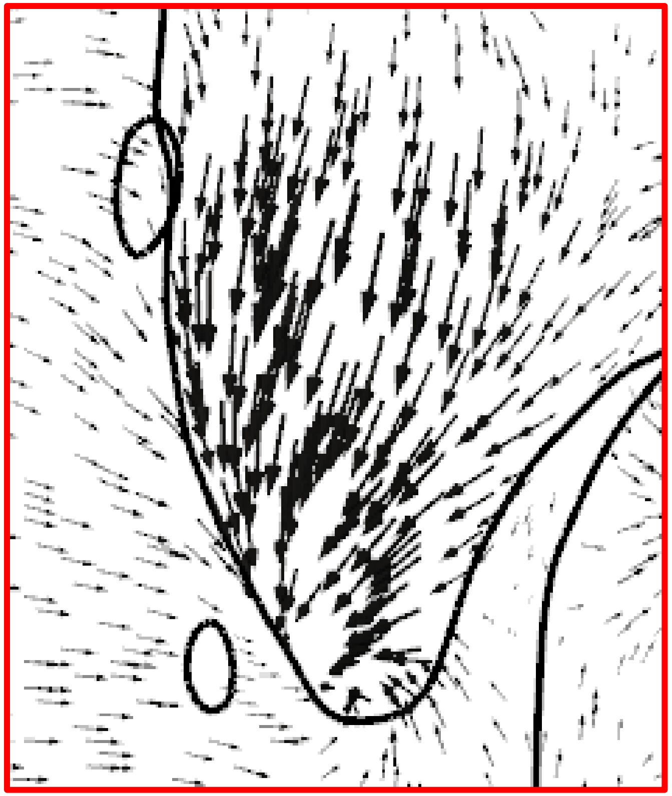
global background motion

pathology grows or contracts



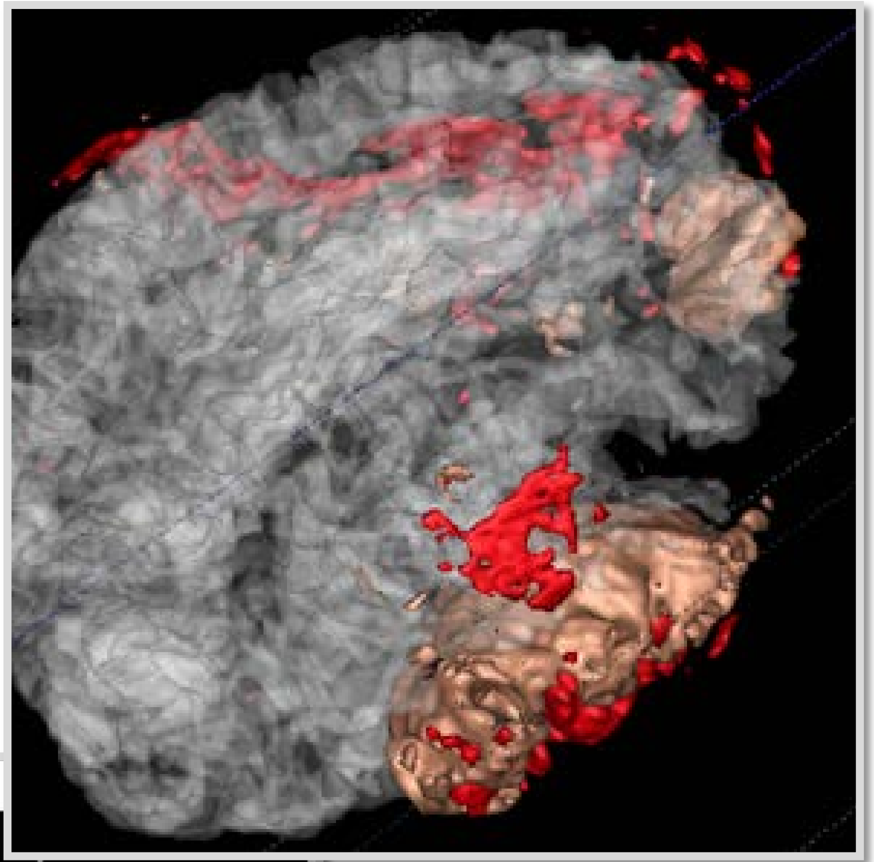
Pace D.F., Niethammer M., Aylward S.R. Sliding Geometries in Deformable Image Registration. Int Conf Med Image Comput Comput Assist Interv. 2011;14(WS):141-148. Workshop on Abdominal Imaging

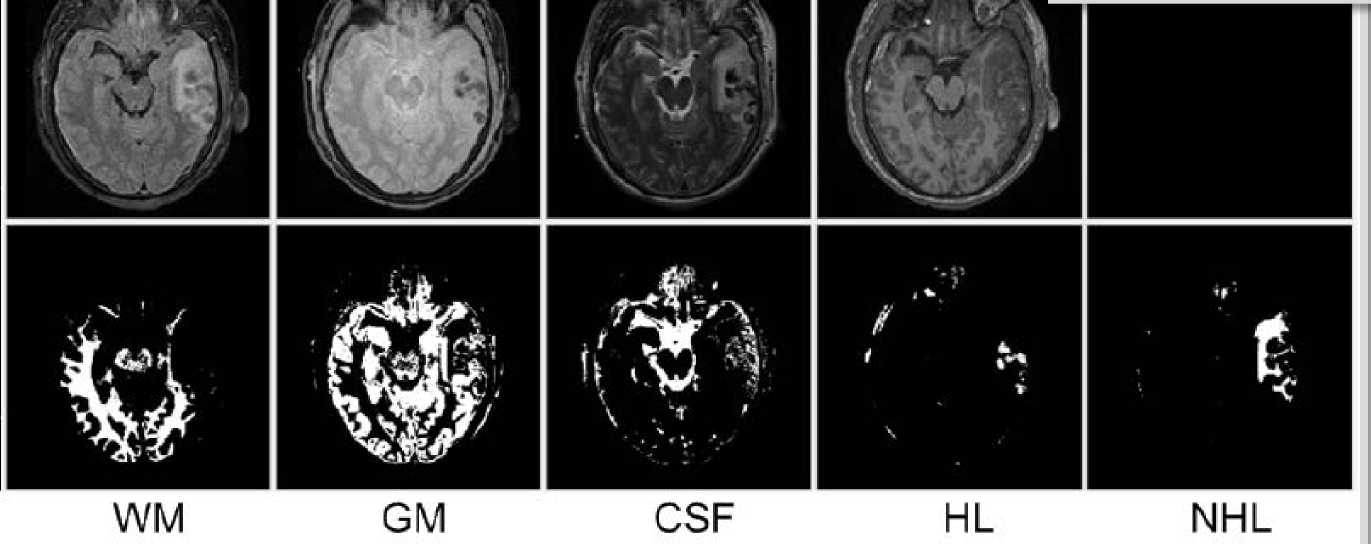
Registration of *sliding* anatomical structures



Algorithm Development 2

Segmentation: Multi-Tissue Type Classification



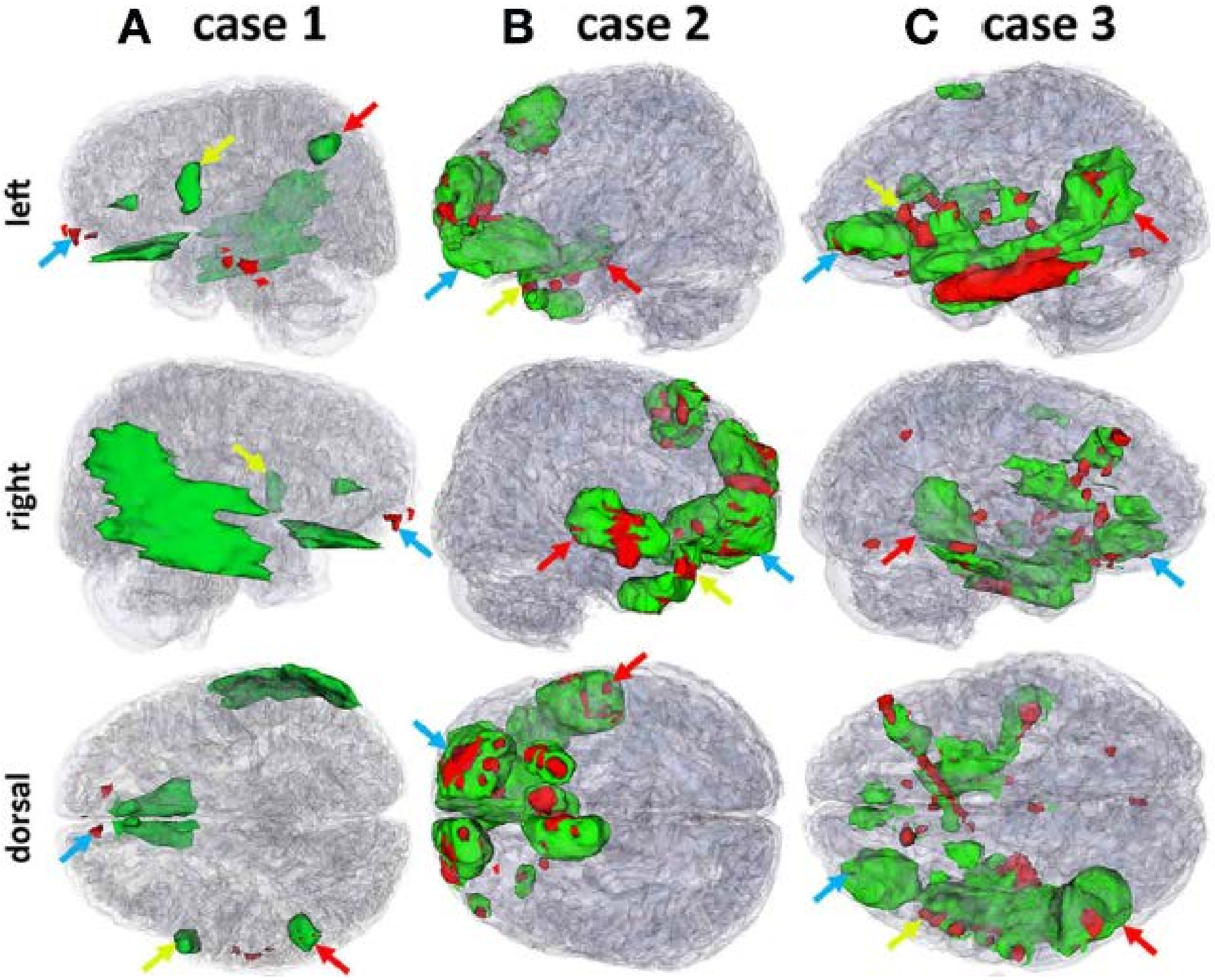


Wang B., Prastawa M., Irimia A., Chambers M.C., Vespa P.M., Van Horn J.D., Gerig G. *A Patient-specific Segmentation Framework for Longitudinal MR Images of Traumatic Brain Injury*. Proceedings of SPIE 2012;8314, 831402.

Subject Specific Analysis 1

3D models of pathology in three patients

Wang B., Prastawa M., Irimia A., Chambers M.C., Vespa P.M., Van Horn J.D., Gerig G. *A Patient-specific Segmentation Framework for Longitudinal MR Images of Traumatic Brain Injury*. Proceedings of SPIE 2012;8314, 831402.



🟩 edema

🟥 hemorrhage

Subject Specific Analysis 2

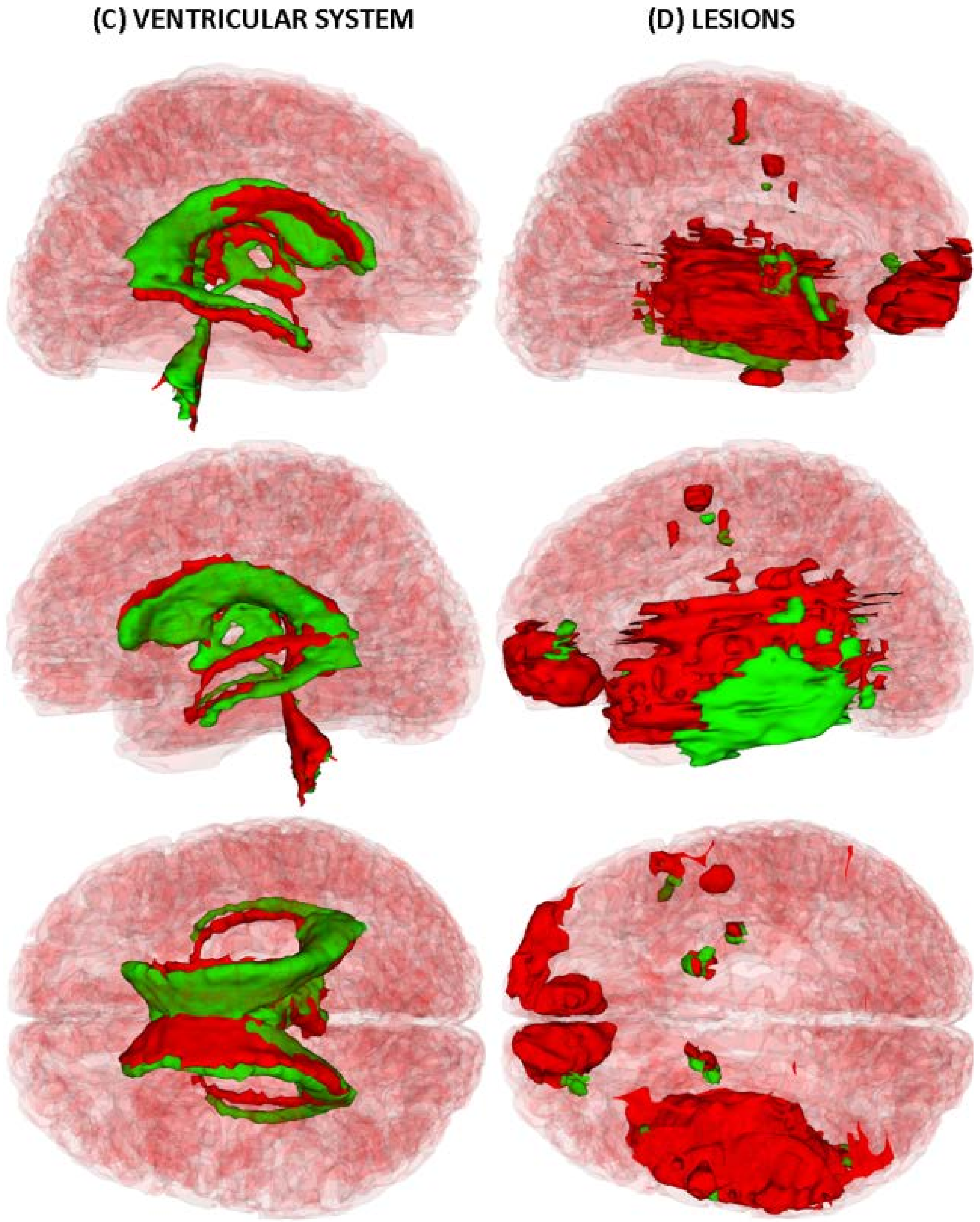
Longitudinal change:
acute vs. chronic

Wang B., Prastawa M., Irimia A., Chambers M.C., Vespa P.M., Van Horn J.D., Gerig G. *A Patient-specific Segmentation Framework for Longitudinal MR Images of Traumatic Brain Injury*. Proceedings of SPIE 2012;8314, 831402.

🟥 acute

🟩 chronic

Wang et al., Univ. Utah



Group Comparisons

- Clinical Imaging Studies are similar to the basic science paradigm discussed previously:
  - Large number of subjects
  - Years for the analysis
- Requires fully automated pipelines
- Requires large computational resources

- Example: COPD gene

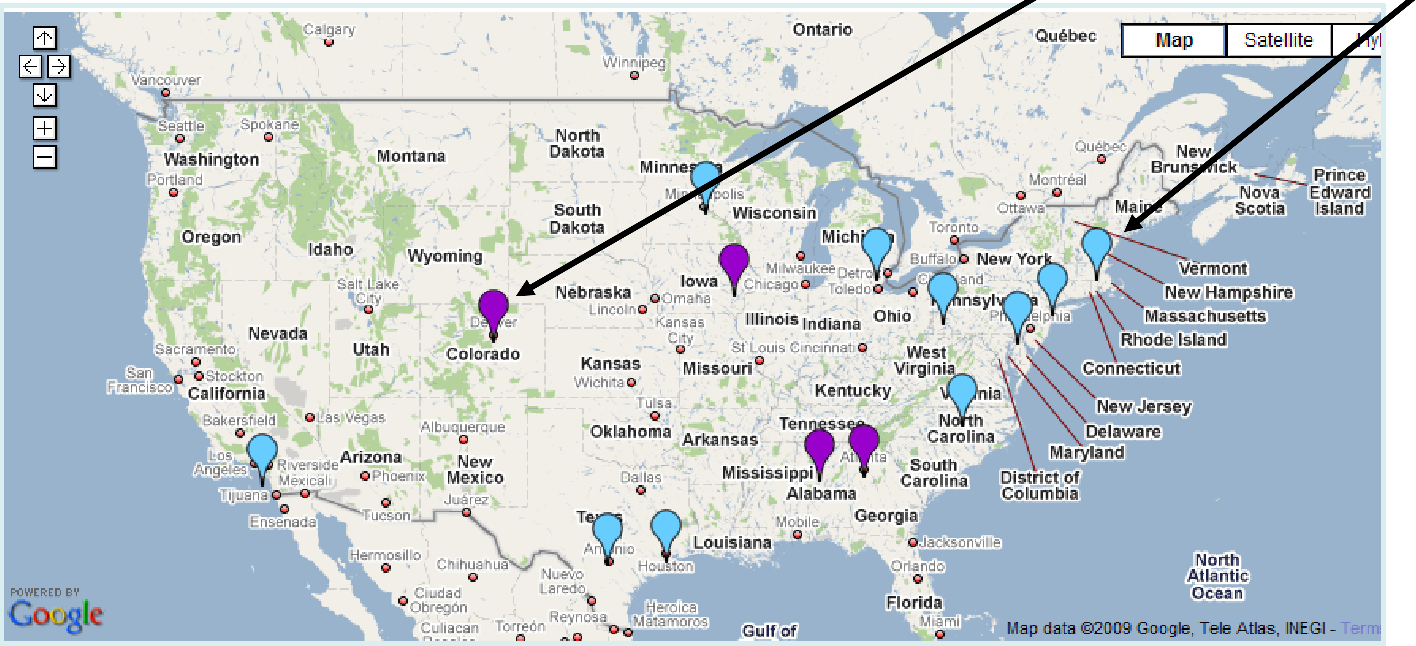
COPDGene

Only 20% of smokers develop COPD → Genetic factors

COPDGene®
COPD Genetic Epidemiology

Multi-center study funded by the National Heart, Lung and Blood Institute (NHLBI).

Co-PIs: Drs. James Crapo, Edwin Silverman.



21 clinical sites

3 image analysis centers:

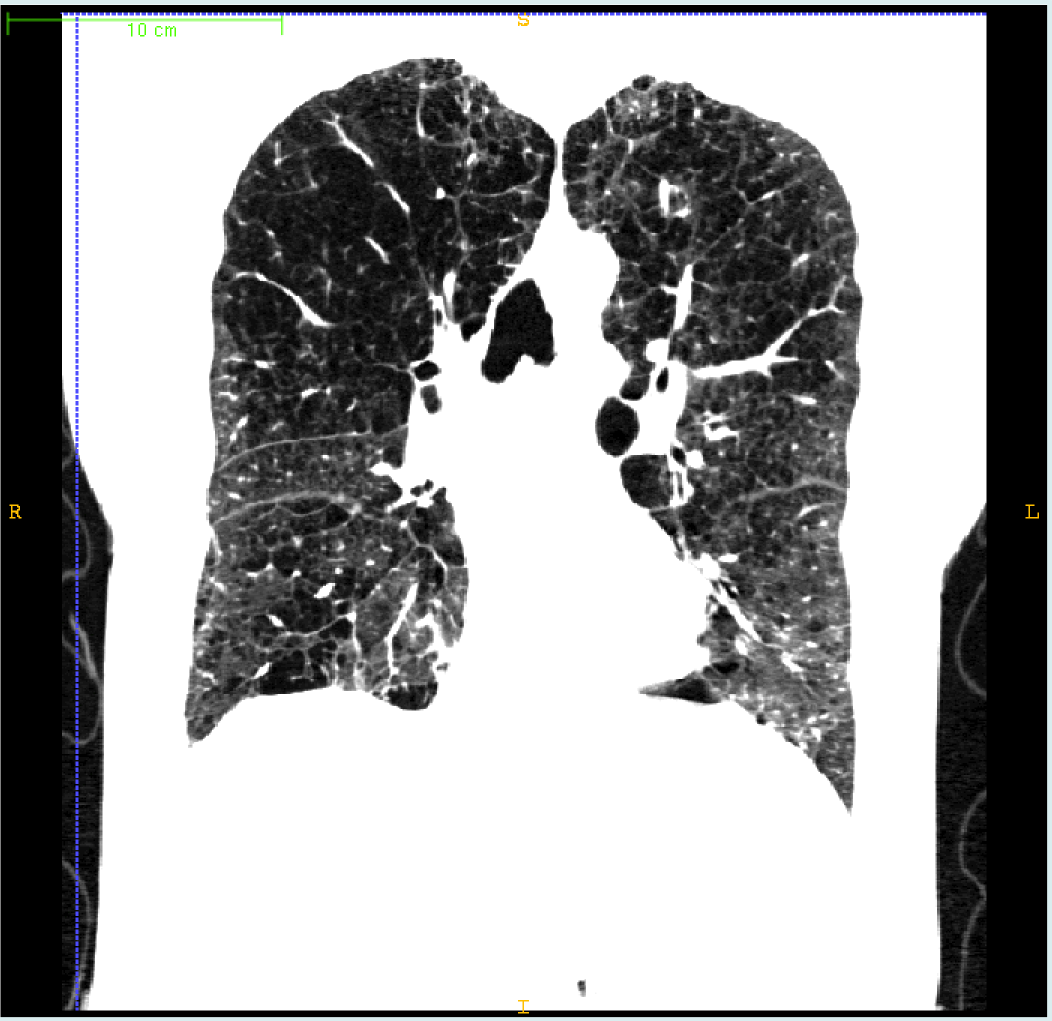
- Denver
- Boston
- Iowa

2 imaging platforms:

- VIDA
- Slicer

Slide courtesy R. San Jose

Emphysema Classification for Gene Discovery



- Identification of emphysema patterns based on local histogram classification

🟥 Normal
🟨 Mild CLE
🟦 Moderate CLE
🟪 Severe CLE
🟦 PLE
🟩 Paraseptal

- Centrilobular (CLE) and panacinar (PLE) emphysema
- GWAS in 9000 smokers
- New genetic markers for emphysema were found near the CHRNA3/5 locus on 15q25 and near MMP12 and MMP3 on 11q22

Castaldi PJ, San Jose Estepar R, Sanchez Mendoza C, Crapo JD, Lynch D, Beaty TH, Washko GR, Silverman EK, Proc. ATS, 2012, p.A3808.

Slide courtesy R. San Jose

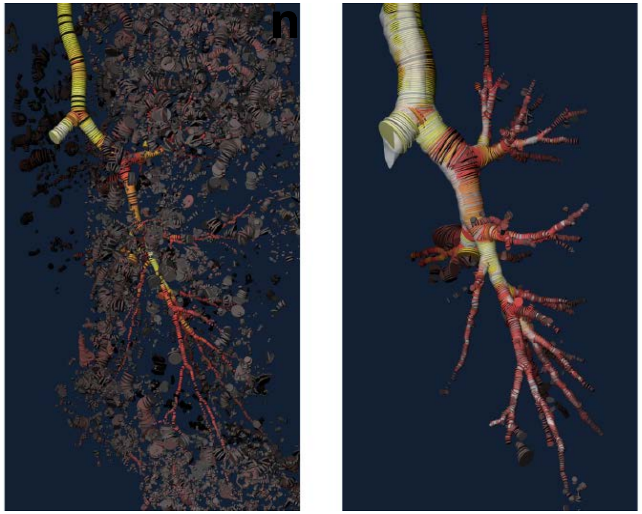
Phenotype Analysis In The Lung

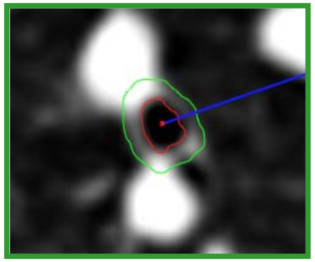
Extraction

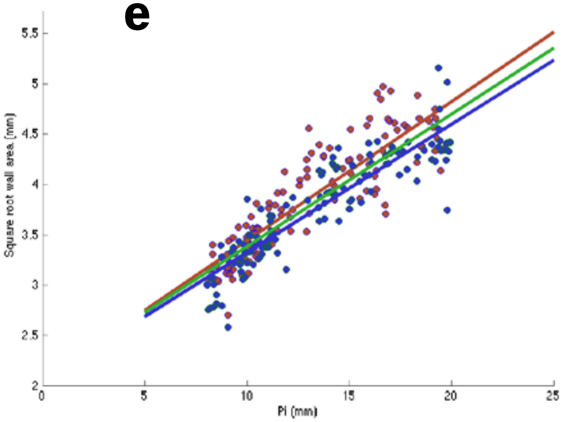
Sizing

Phenotype

Airways

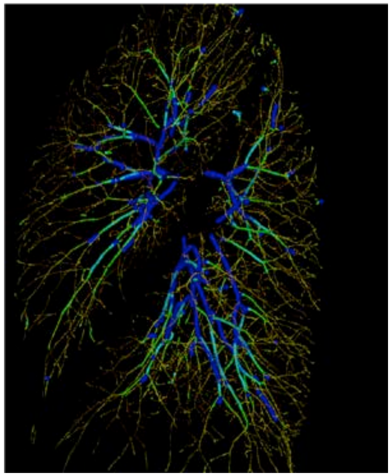




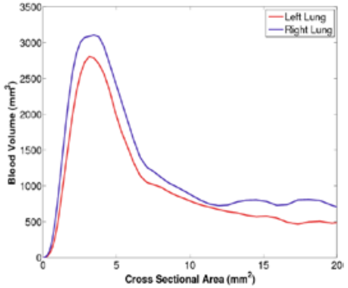


Airway wall corresponding to a 10 mm internal Perimeter

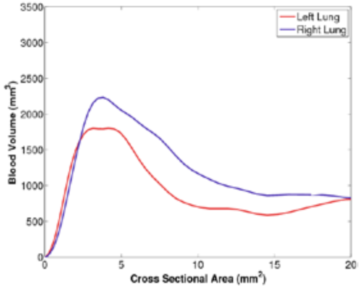
Vessels



Smoker control



Severe disease



Shift in blood volume per cross sectional area indicating distal pruning and proximal remodeling with disease progression

Kindlmann G, San José Estépar R, Smith SM, Westin CF. Sampling and visualizing creases with scale-space particles. IEEE Transactions on Visualization and Computer Graphics 2009;15(6):1415-1424. PMID: 19834216.

THEME C: Biomedical Informatics Under the NCBC Program

Mark A. Musen

Stanford University

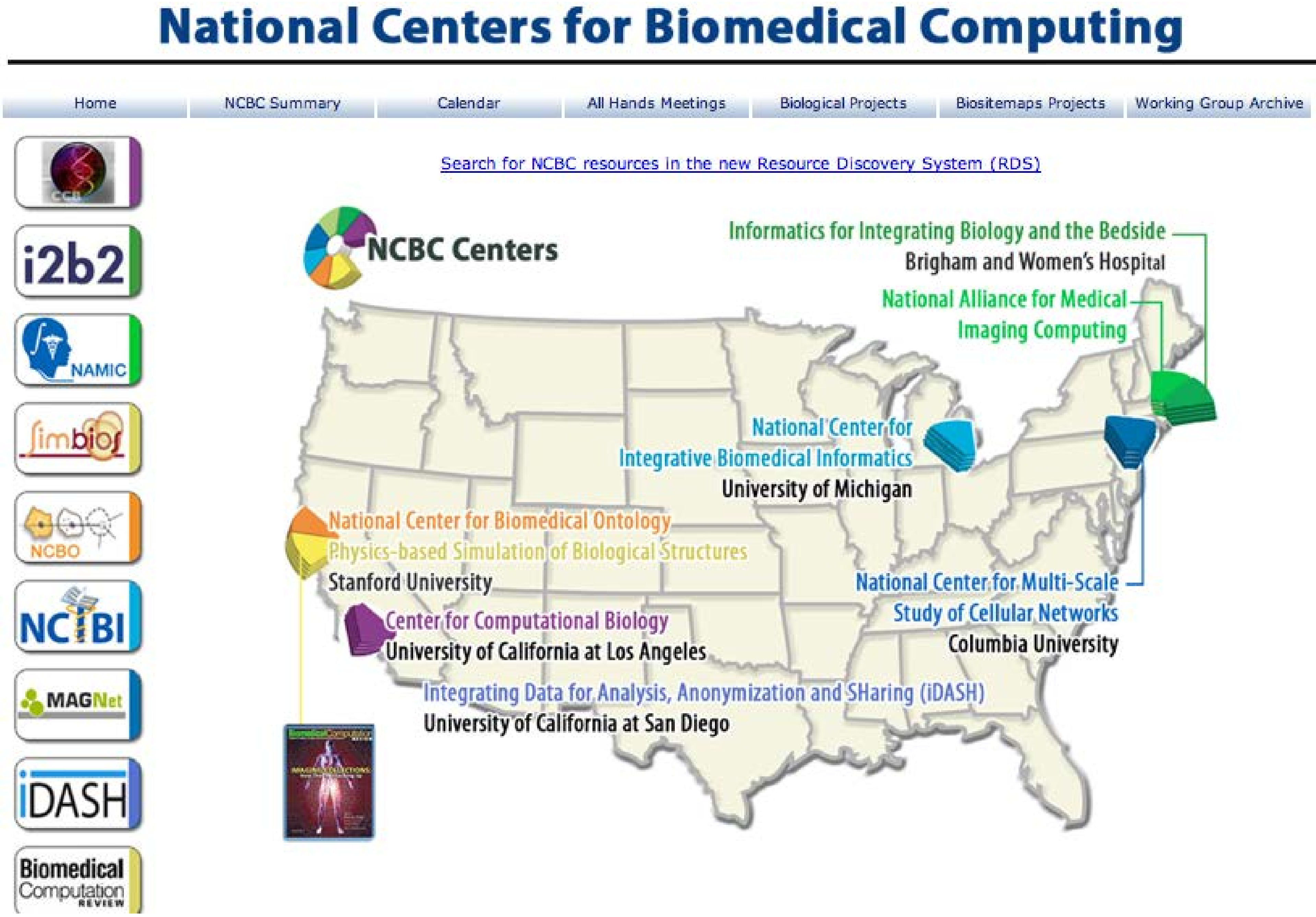
iDASH Integrating Data for Analysis, Anonymization, and SHaring

i2b2

Informatics for Integrating Biology & the Bedside

NATIONAL CENTER FOR

BIOMEDICAL ONTOLOGY



National Centers for Biomedical Computing

Search for NCBC resources in the new Resource Discovery System (RDS)

NCBC Centers

Informatics for Integrating Biology and the Bedside
Brigham and Women's Hospital

National Alliance for Medical Imaging Computing

National Center for Integrative Biomedical Informatics
University of Michigan

National Center for Biomedical Ontology
Physics-based Simulation of Biological Structures
Stanford University

National Center for Multi-Scale Study of Cellular Networks
Columbia University

Center for Computational Biology
University of California at Los Angeles

Integrating Data for Analysis, Anonymization and SHaring (iDASH)
University of California at San Diego

Bioinformatics in the NCBC Program

- Focus is on infrastructure for
  - Data modeling
  - Data acquisition
  - Data management
  - Data analysis
- Continuum of data from basic biology to clinical medicine to population health
- Each of the centers in this group emphasizes different methodological aspects

Bioinformatics-related NCBCs

i2b2

Informatics for Integrating Biology & the Bedside

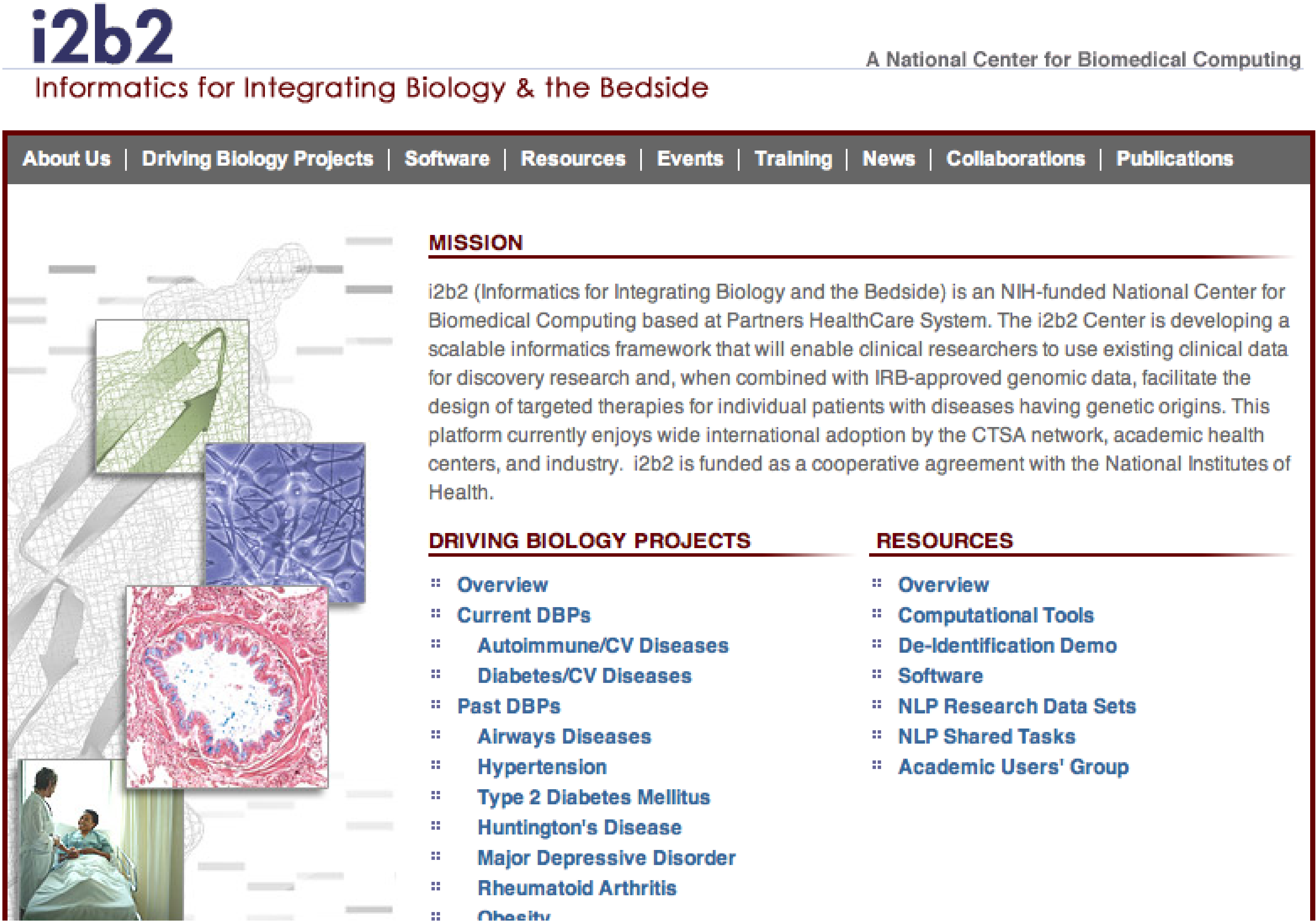
Isaac Kohane, PI; Harvard/Partners' HealthCare

iDASH

Lucila Ohno-Machado, PI; UCSD

NATIONAL CENTER FOR
BIOMEDICAL ONTOLOGY

Mark Musen, PI; Stanford



i2b2

Informatics for Integrating Biology & the Bedside

A National Center for Biomedical Computing

MISSION

i2b2 (Informatics for Integrating Biology and the Bedside) is an NIH-funded National Center for Biomedical Computing based at Partners HealthCare System. The i2b2 Center is developing a scalable informatics framework that will enable clinical researchers to use existing clinical data for discovery research and, when combined with IRB-approved genomic data, facilitate the design of targeted therapies for individual patients with diseases having genetic origins. This platform currently enjoys wide international adoption by the CTSA network, academic health centers, and industry. i2b2 is funded as a cooperative agreement with the National Institutes of Health.

DRIVING BIOLOGY PROJECTS

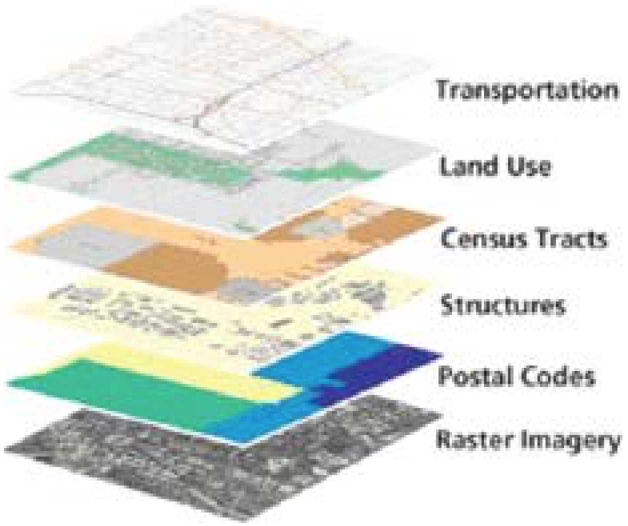
- Overview
- Current DBPs
  - Autoimmune/CV Diseases
  - Diabetes/CV Diseases
- Past DBPs
  - Airways Diseases
  - Hypertension
  - Type 2 Diabetes Mellitus
  - Huntington's Disease
  - Major Depressive Disorder
  - Rheumatoid Arthritis
  - Obesity

RESOURCES

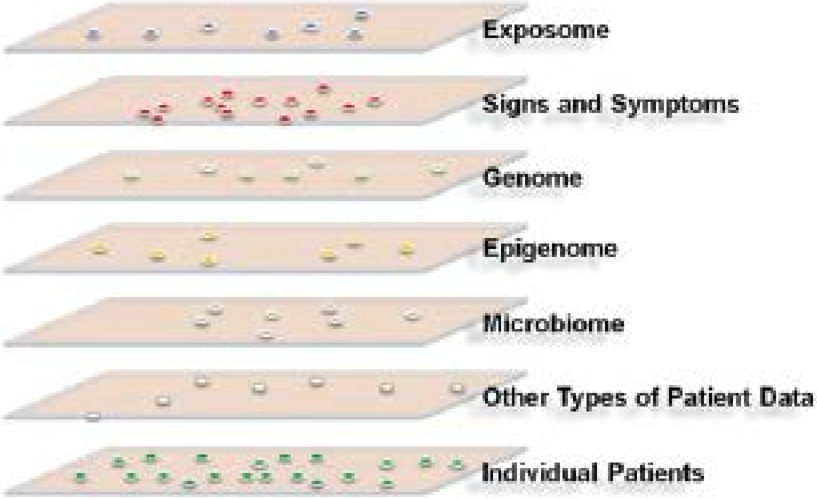
- Overview
- Computational Tools
- De-Identification Demo
- Software
- NLP Research Data Sets
- NLP Shared Tasks
- Academic Users' Group

Better, faster insights through for Precision Medicine by Instrumenting Health care and Populations

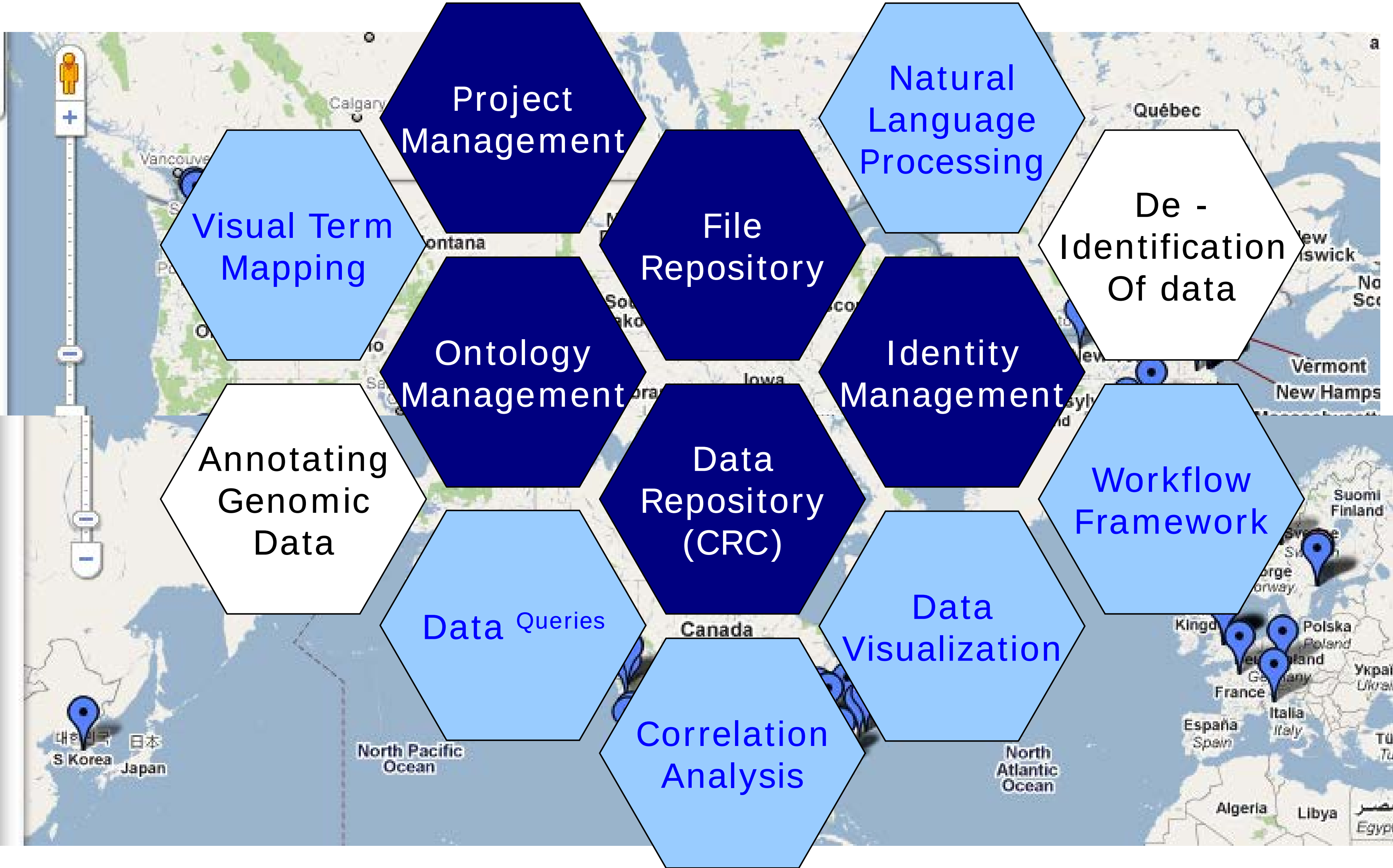
Google Maps: GIS layers
Organized by Geographical Positioning



Information Commons
Organized Around Individual Patients



National Translational Toolkit



Project Management

Natural Language Processing

Visual Term Mapping

File Repository

De - Identification Of data

Ontology Management

Identity Management

Annotating Genomic Data

Data Repository (CRC)

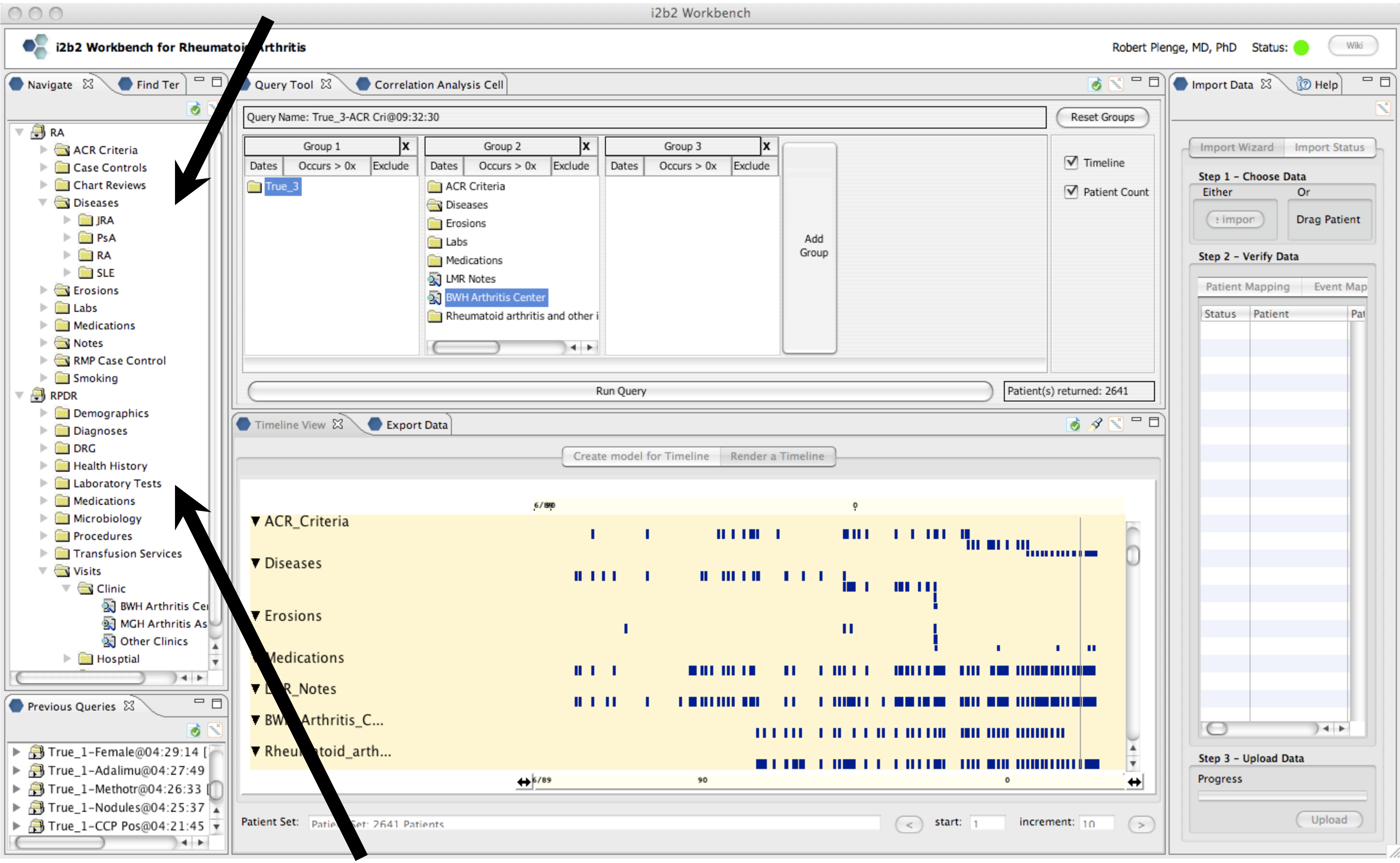
Workflow Framework

Data Queries

Data Visualization

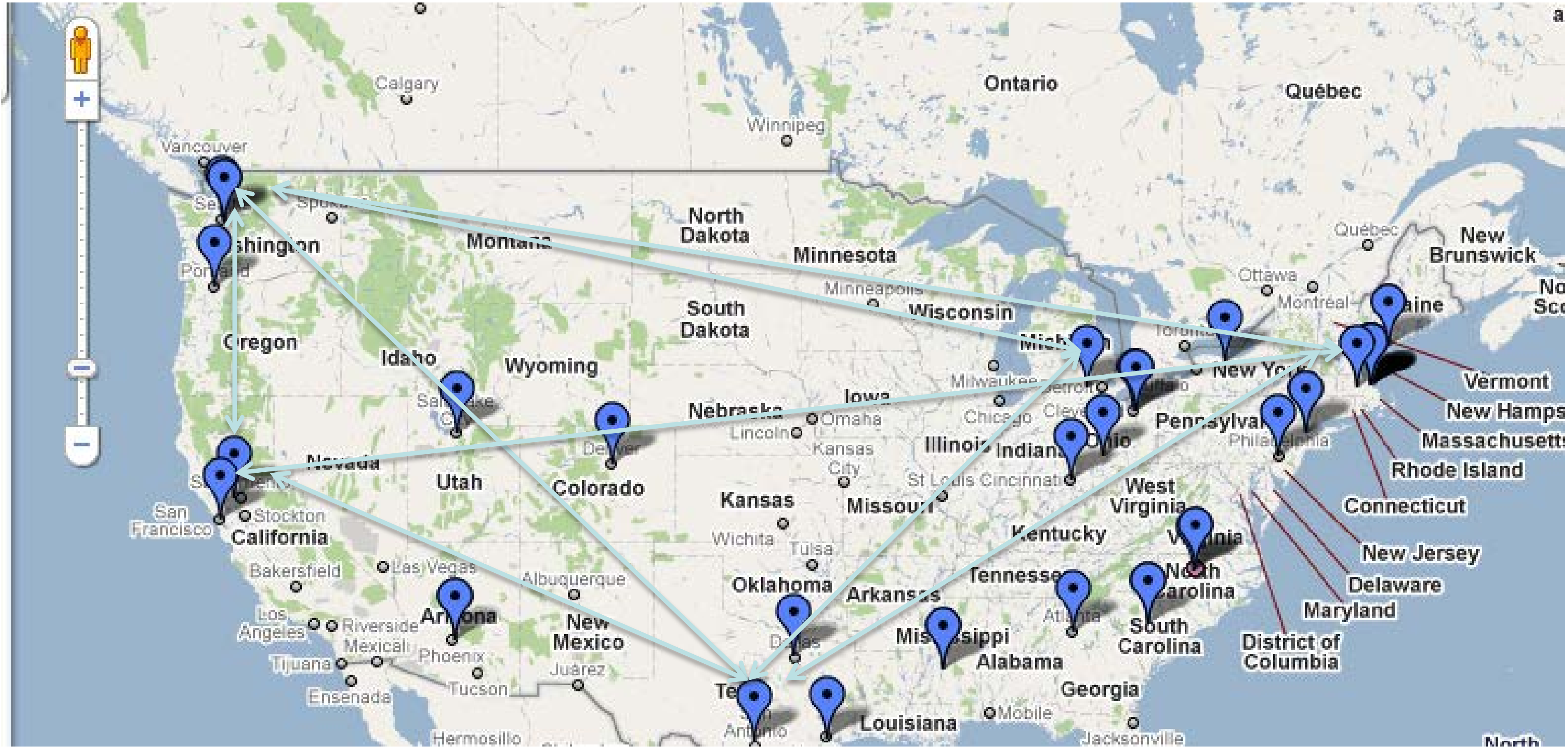
Correlation Analysis

Narrative data (NLP text extractions)



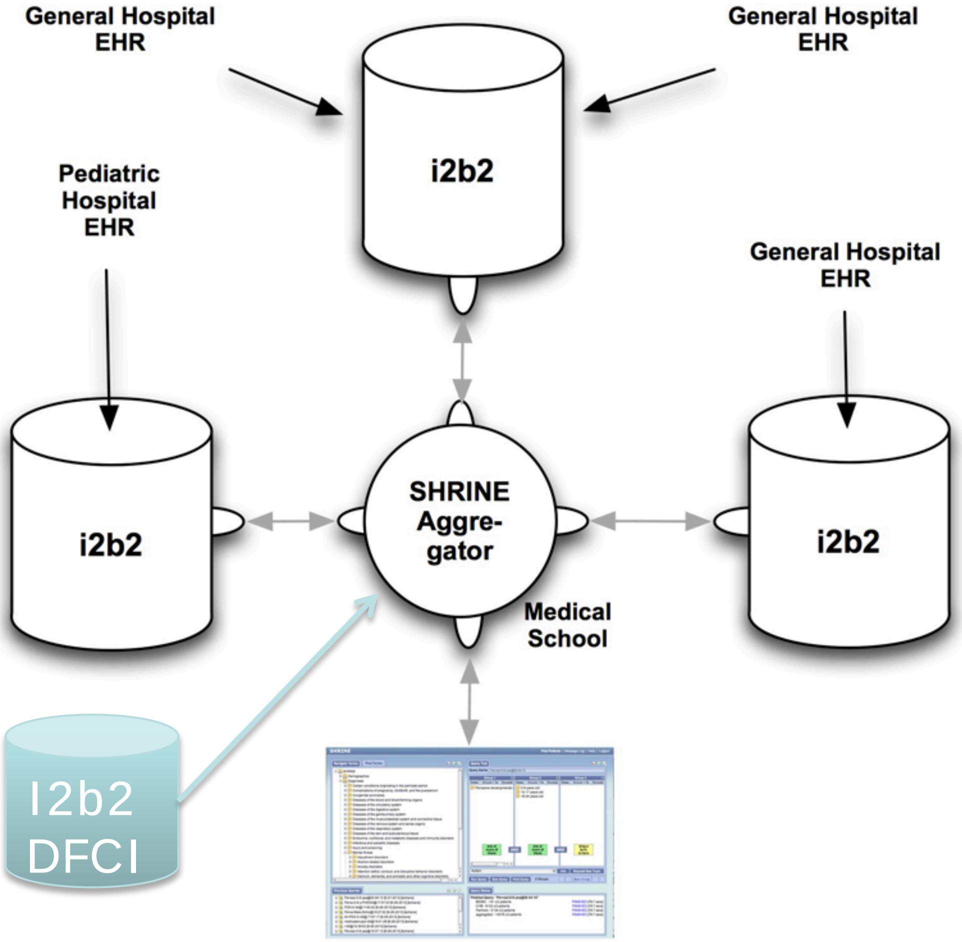
Codified data (ICD9 codes, etc)

2012

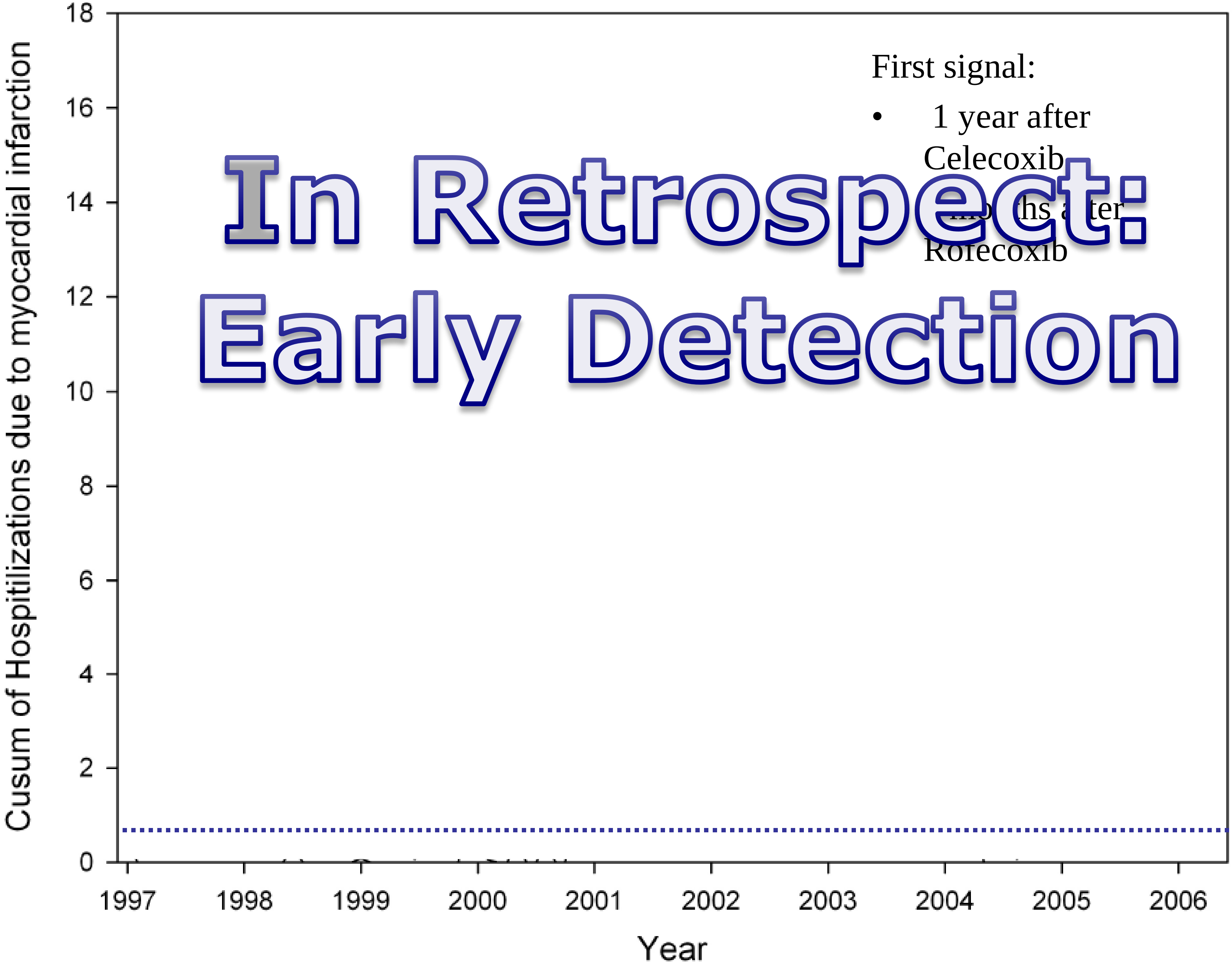

Calgary
Winnipeg
Ontario
Québec
Vancouver
Spokane
Washington
Portland
Montana
North Dakota
Minnesota
South Dakota
Wisconsin
Minneapolis
Ottawa
Montréal
Québec
New Brunswick
Maine
Oregon
Idaho
Wyoming
Michigan
Toronto
New York
Vermont
New Hampshire
Massachusetts
Rhode Island
Connecticut
Salt Lake City
Nebraska
Omaha
Lincoln
Iowa
Milwaukee
Detroit
Chicago
Cleveland
Buffalo
Pennsylvania
Philadelphia
Nevada
Utah
Denver
Colorado
Kansas
Kansas City
Illinois
Indiana
Ohio
St Louis
Cincinnati
West Virginia
San Francisco
Stockton
California
Missouri
Wichita
Tulsa
Kentucky
Virginia
New Jersey
Delaware
Maryland
District of Columbia
Bakersfield
Las Vegas
Albuquerque
Oklahoma
Arkansas
Tennessee
North Carolina
South Carolina
Atlanta
Los Angeles
Riverside
Mexicali
Tijuana
Ensenada
Arizona
Phoenix
Tucson
New Mexico
Juárez
Dallas
Texas
San Antonio
Mississippi
Alabama
Georgia
Louisiana
Mobile
Jacksonville
Hermosillo

SHRINE Network extends the reach of i2b2 across sites

- Search routine clinical records from 5 major hospitals for:
  - Demographics
  - Diagnosis
  - Medications
  - Lab Results
- Reach N
  - Rare Dx
  - Small Effects



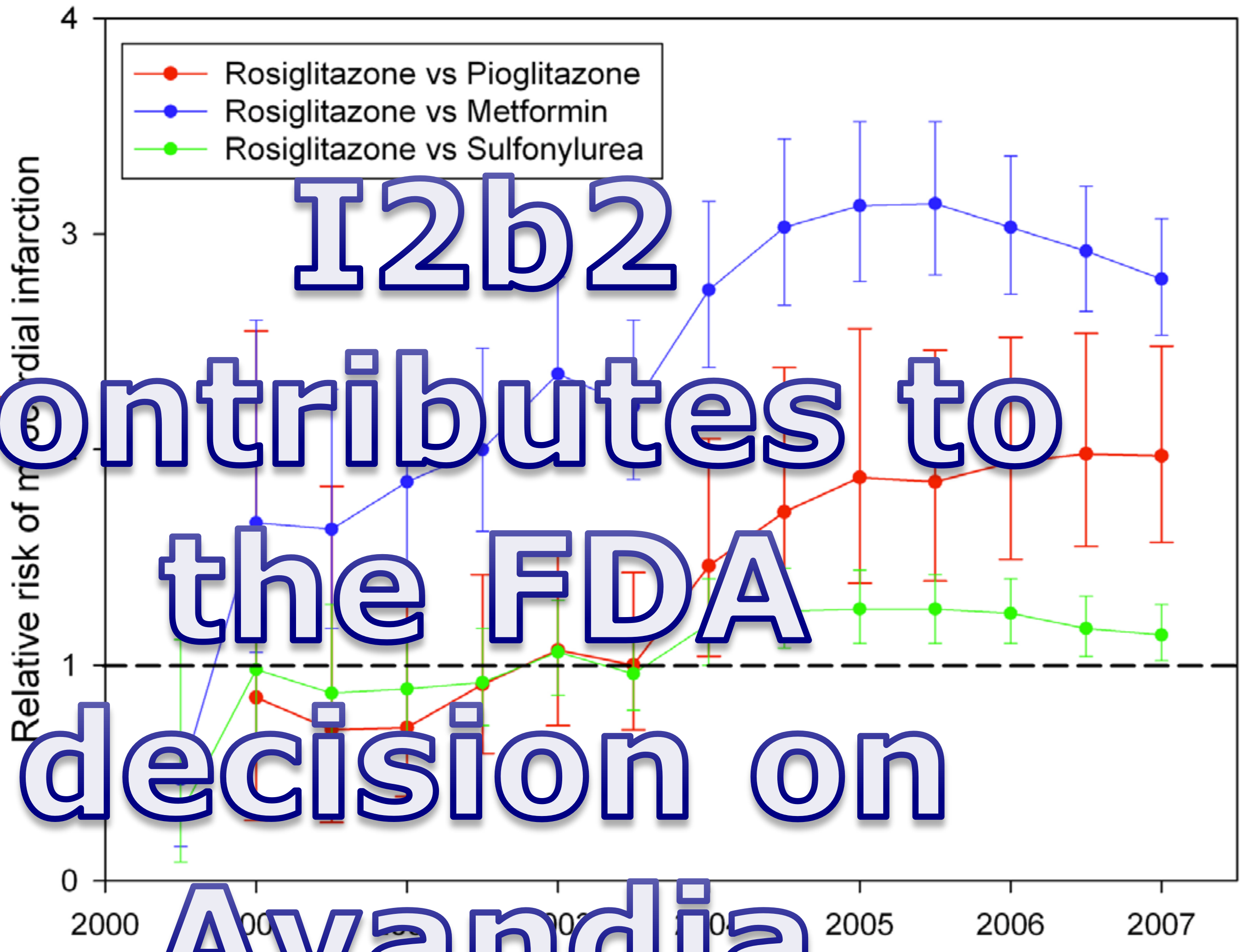
In Retrospect: Early Detection



First signal:

- 1 year after Celecoxib
- [illegible] months after Rofecoxib

Unhidden secrets in our health system



I2b2 contributes to the FDA decision on Avandia

Full Text (PDF)

Rapid Identification of Myocardial Infarction Risk Associated With Diabetes Medications Using Electronic Medical Records

Diabetes Care March 2010 33:526–531; published ahead of print December 15, 2009,

Policy-based Secure Data Sharing

Science
Learn more about our driving biomedical and bioinformatics projects

Share/Access Data
Host and access data in a secure and privacy-preserving environment

Software/Tools
Advance your research and push your boundaries

CyberInfrastructure
Explore the fabric that binds all iDASH resources together

Welcome to iDASH

Contemporary biomedical and behavioral research increasingly demands access to data-intensive computational resources. While the United States' investments in information technology assets has created a rich national fabric to accelerate research, the number of bioscience researchers without these tools is increasing. By integrating secure, patient-anonymous data for analysis, anonymization, and sharing, iDASH is narrowing the gap by extending its infrastructure, tools, and services more broadly to the national biomedical, clinical, and informatics communities at universities, medical schools, and hospitals.

Created as a National Center for Biomedical Computing (NCBC) under the auspices of the NIH Roadmap for Bioinformatics and Computational Biology, iDASH is a Biomedical Cyberinfrastructure (BCI) that provides innovative services, algorithms, open-source software, and data storage as well as training.

iDASH News

iDASH Second Annual All-Hands Symposium
July 18-19, 2012, UCSD, La Jolla, CA

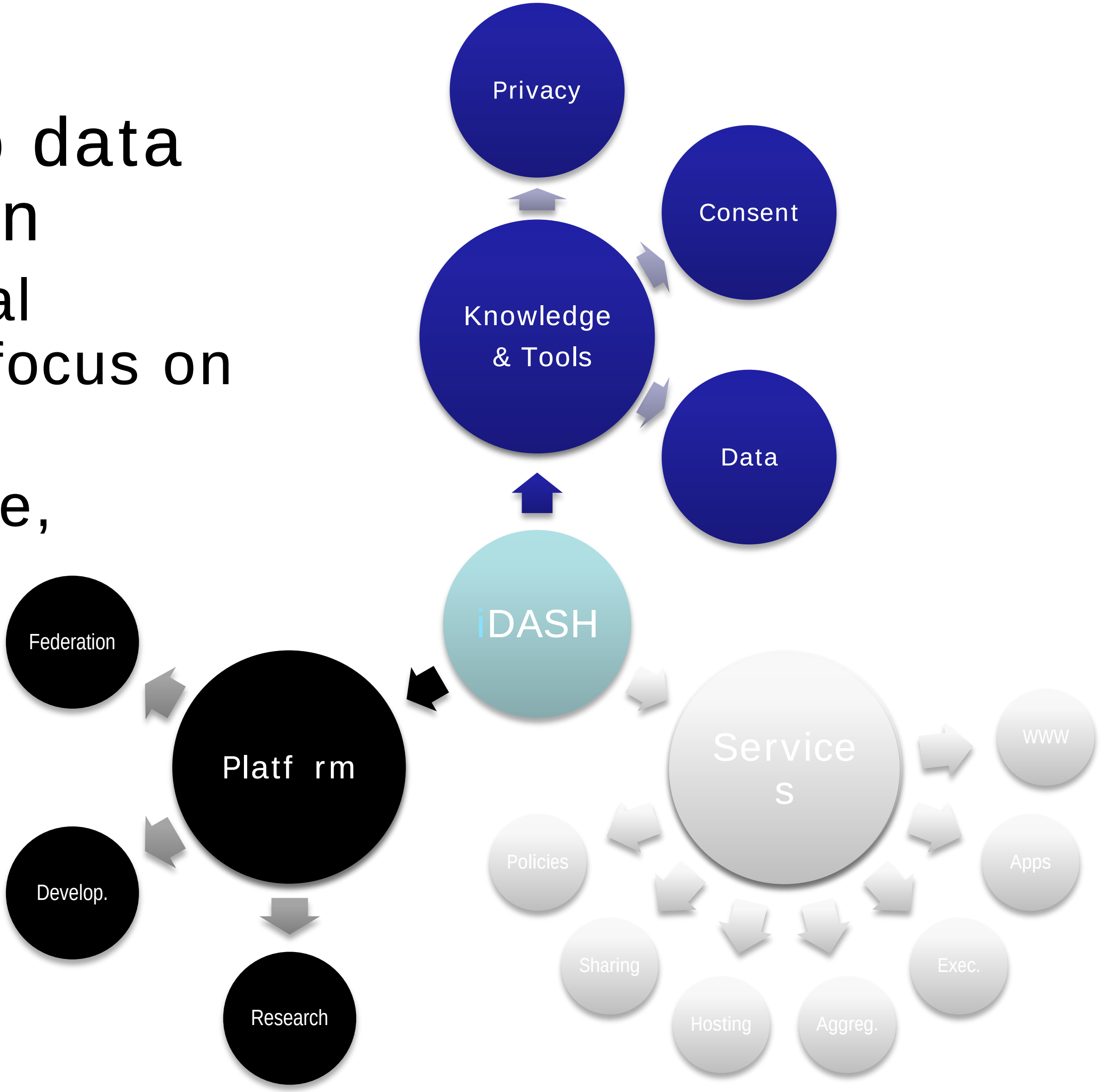
Biomedical Data Sharing: Ethical, Legal, and Policy Perspectives
July 18, 2012, UCSD, La Jolla, CA
Details

IEEE Conference on Healthcare Informatics, Imaging, and Systems Biology (HISB)
September 27-28, 2012, UCSD, La Jolla, CA
Details

iDASH Privacy Workshop
September 29, 2012, UCSD, La Jolla, CA

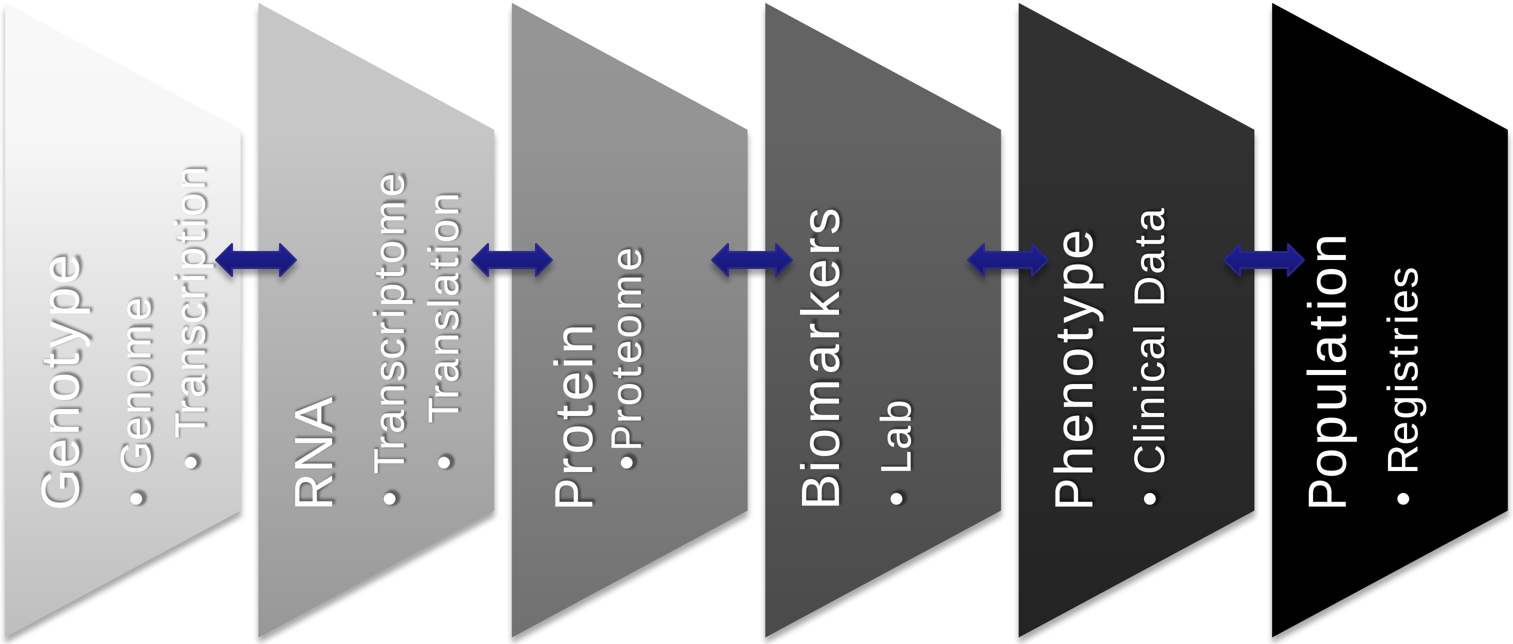
iDASH Vision

- Share access to data and computation
  - Allow biomedical researchers to focus on their science
  - Provide software, platform, and infrastructure
  - Protect privacy
  - Share
    - Data
    - Workflows
    - Computation
    - Security
    - Policies



Privacy
Consent
Knowledge & Tools
Data
iDASH
Federation
Platform
Develop.
Research
Services
WWW
Apps
Policies
Sharing
Hosting
Aggreg.
Exec.

Big Data, Small Data, and other Data



Genotype
- Genome
- Transcription

RNA
- Transcriptome
- Translation

Protein
- Proteome

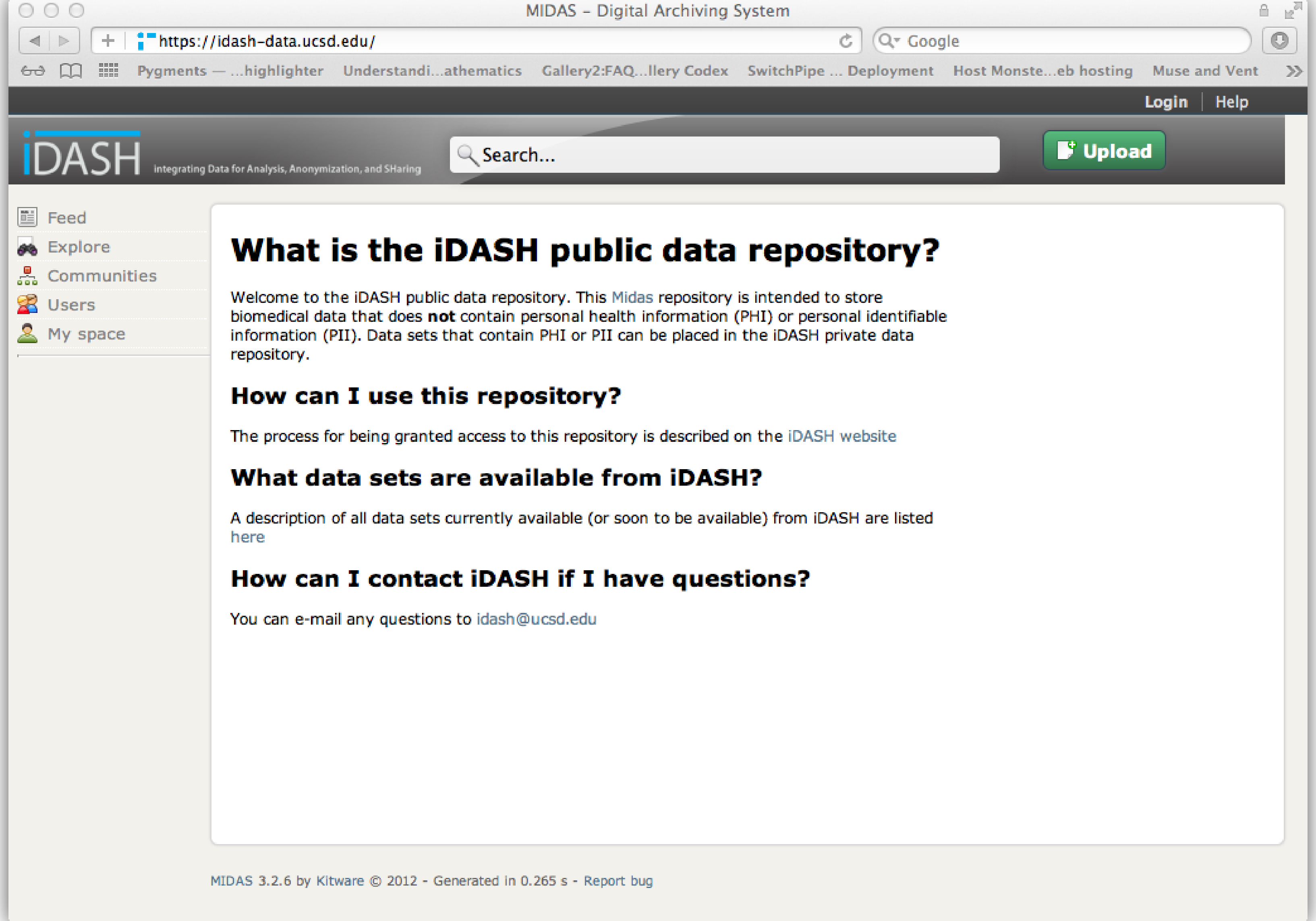
Biomarkers
- Lab

Phenotype
- Clinical Data

Population
- Registries

- Data integration across biological scales
- Data analysis from multiple sources
- Data ‘anonymization’ and privacy preservation

Public Repository for Biomedical Data



What is the iDASH public data repository?

Welcome to the iDASH public data repository. This Midas repository is intended to store biomedical data that does **not** contain personal health information (PHI) or personal identifiable information (PII). Data sets that contain PHI or PII can be placed in the iDASH private data repository.

How can I use this repository?

The process for being granted access to this repository is described on the iDASH website

What data sets are available from iDASH?

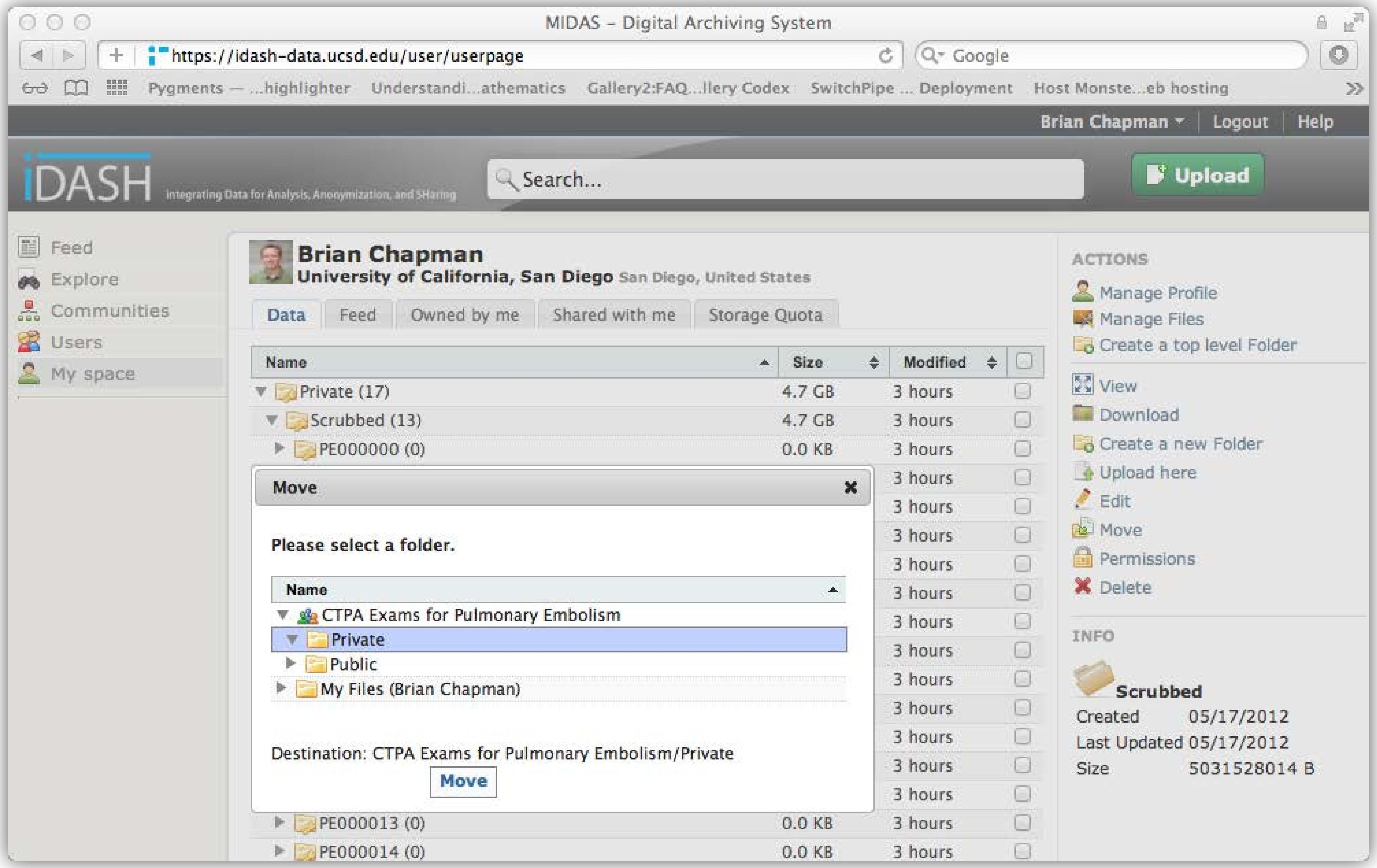
A description of all data sets currently available (or soon to be available) from iDASH are listed here

How can I contact iDASH if I have questions?

You can e-mail any questions to idash@ucsd.edu

MIDAS 3.2.6 by Kitware © 2012 - Generated in 0.265 s - Report bug

Private Repository for Biomedical Data



Brian Chapman
University of California, San Diego San Diego, United States

Name	Size	Modified	☐
▼ Private (17)	4.7 GB	3 hours	☐
▼ Scrubbed (13)	4.7 GB	3 hours	☐
▶ PE000000 (0)	0.0 KB	3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
		3 hours	☐
▶ PE000013 (0)	0.0 KB	3 hours	☐
▶ PE000014 (0)	0.0 KB	3 hours	☐

Move ✖

Please select a folder.

Name
▼ CTPA Exams for Pulmonary Embolism
▼ Private
▶ Public
▶ My Files (Brian Chapman)

Destination: CTPA Exams for Pulmonary Embolism/Private

Move

INFO

Scrubbed

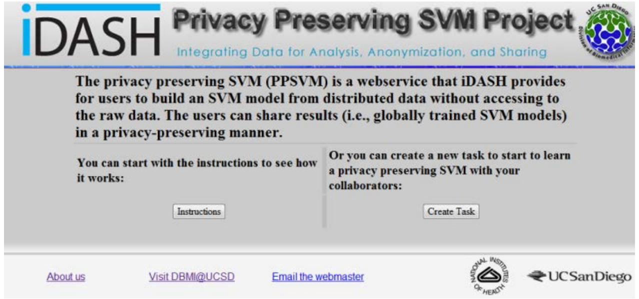
Created	05/17/2012
Last Updated	05/17/2012
Size	5031528014 B

Models for Sharing

- *Traditional*: iDASH-hosted data exported for computation
  - Users download data from iDASH
  - Good for Small Data, Low-risk data with Data Use Agreements
- Computation comes to iDASH-data
  - (Users download artificial sets to pre-test their algorithms)
  - Users upload algorithms into iDASH
  - Good for Big Data, High Performance Computing, Genomics, PHI
- iDASH-computation goes to the data
  - Distributed computation, shared modeling
  - Good for when data cannot leave the institution/country

Privacy preserving model building

- Distributed Privacy Preserving Support Vector machine (DPP-SVM)
  - http://privacy.ucsd.edu:8080/ppsvm
  - Ready for use



The privacy preserving SVM (PPSVM) is a webservice that iDASH provides for users to build an SVM model from distributed data without accessing to the raw data. The users can share results (i.e., globally trained SVM models) in a privacy-preserving manner.

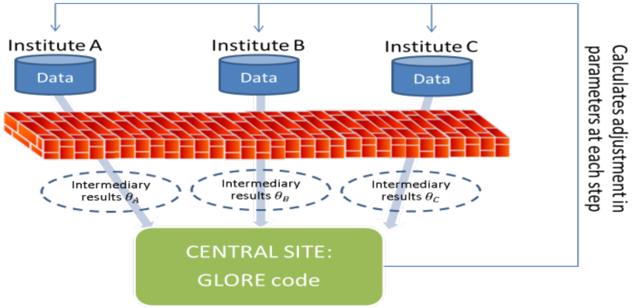
You can start with the instructions to see how it works:

Or you can create a new task to start to learn a privacy preserving SVM with your collaborators:

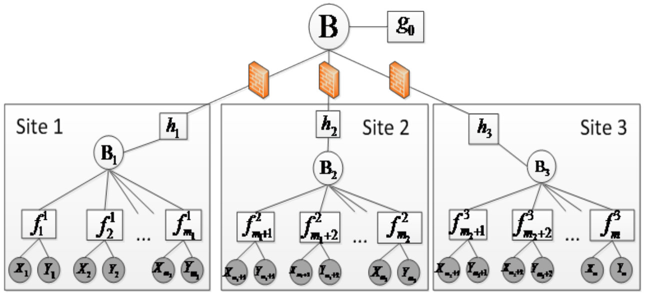
- Grid LOgistic REgression (GLORE)
  - Demo: http://tinyurl.com/7fwnkf6

  UI in development (3 months), CLI ready

  *Wu Y, et al. JAMIA 2012



- EXpectation Propagation based LOgistic REgRession (EXPLORER)
  - Demo: http://tinyurl.com/7b5xaab
  - UI in development (3 months) , CLI ready



Of Current Interest

- **News**: Join the NCBO team at SemTech 2012 and get a "Friends of NCBO" discount
- **Webinar**: Suggest a Speaker
- **Recent Publication**: Whetzel, PL, et al. (2011): BioPortal: enhanced functionality via new Web services...
- **Recent Release**: BioPortal 3.9 (June 2012)
- **NCBO Webinar Announcements - Subscribe**
- **NCBO Software Support - Mailing List Archive**
- More News & Events

Follow us on Twitter

NCBO User Profile

Jin-Dong Kim
Database Center for Life Science

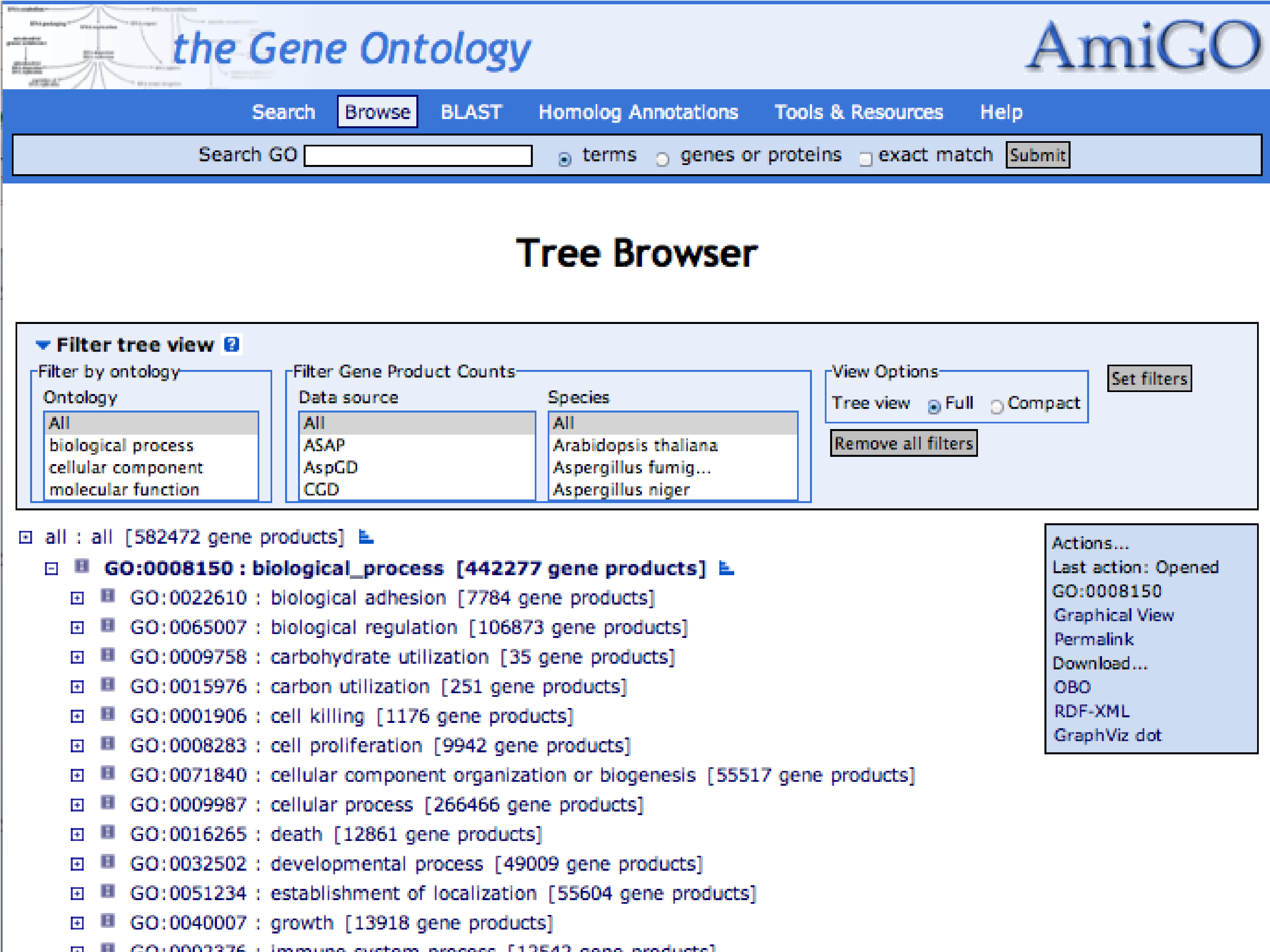
More >
Other profiles >

Video

Learn about Biomedical Ontologies. Watch a series of introductory videos.

Browse ontologies in BioPortal!

BioPortal allows users to browse, search and visualize ontologies.



Tree Browser

Filter tree view

Filter by ontology
Ontology: All | biological process | cellular component | molecular function

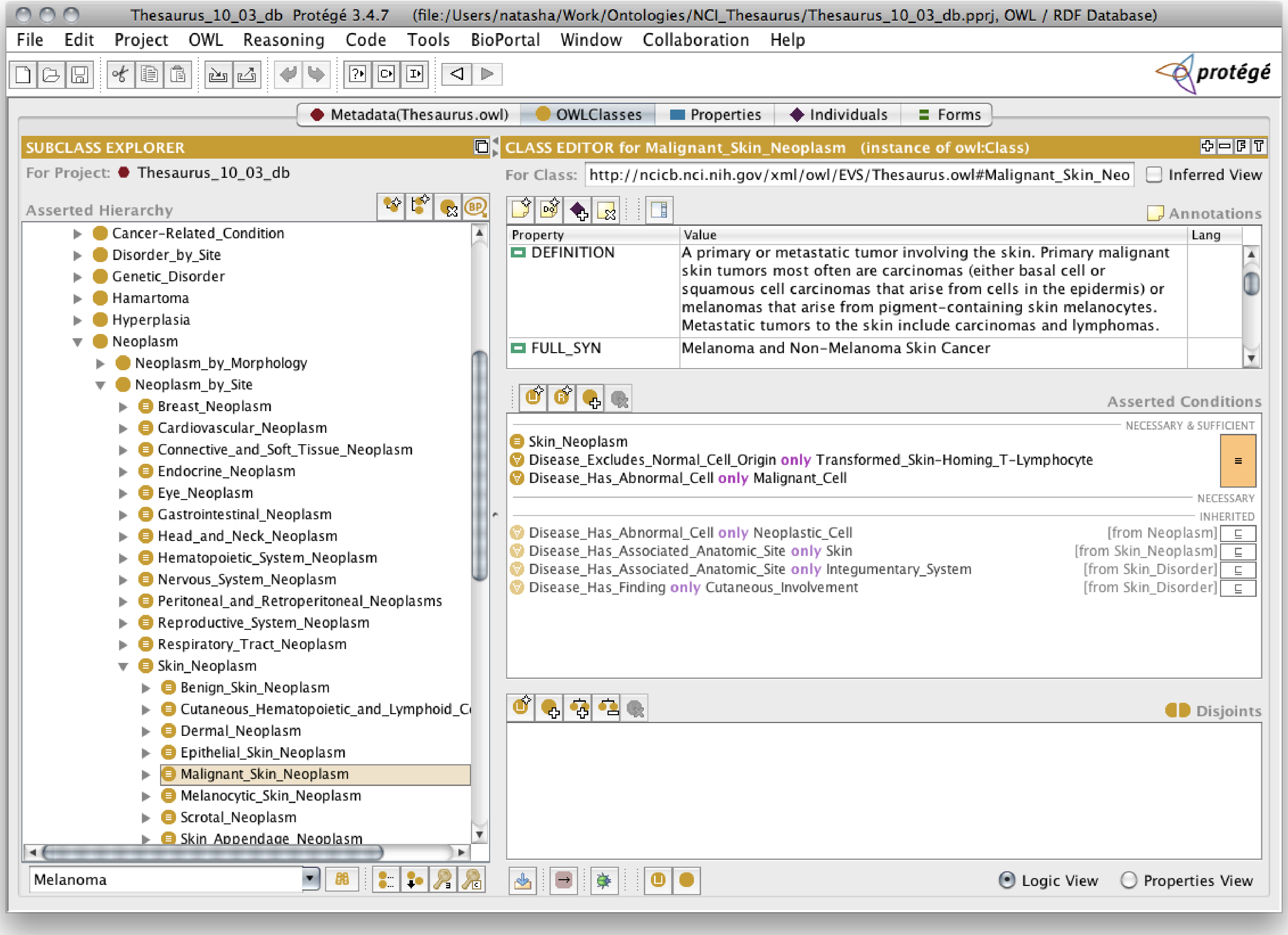
Filter Gene Product Counts
Data source: All | ASAP | AspGD | CGD
Species: All | Arabidopsis thaliana | Aspergillus fumig... | Aspergillus niger

View Options
Tree view ● Full ○ Compact

⊞ all : all [582472 gene products]
- ⊟ **GO:0008150 : biological_process [442277 gene products]**
  - ⊞ GO:0022610 : biological adhesion [7784 gene products]
  - ⊞ GO:0065007 : biological regulation [106873 gene products]
  - ⊞ GO:0009758 : carbohydrate utilization [35 gene products]
  - ⊞ GO:0015976 : carbon utilization [251 gene products]
  - ⊞ GO:0001906 : cell killing [1176 gene products]
  - ⊞ GO:0008283 : cell proliferation [9942 gene products]
  - ⊞ GO:0071840 : cellular component organization or biogenesis [55517 gene products]
  - ⊞ GO:0009987 : cellular process [266466 gene products]
  - ⊞ GO:0016265 : death [12861 gene products]
  - ⊞ GO:0032502 : developmental process [49009 gene products]
  - ⊞ GO:0051234 : establishment of localization [55604 gene products]
  - ⊞ GO:0040007 : growth [13918 gene products]
  - ⊞ GO:0002376 : immune system process [12542 gene products]

Actions...
Last action: Opened GO:0008150
Graphical View
Permalink
Download...
OBO
RDF-XML
GraphViz dot

The National Cancer Institute Thesaurus



For Project: Thesaurus_10_03_db

- Cancer-Related_Condition
- Disorder_by_Site
- Genetic_Disorder
- Hamartoma
- Hyperplasia
- Neoplasm
  - Neoplasm_by_Morphology
  - Neoplasm_by_Site
    - Breast_Neoplasm
    - Cardiovascular_Neoplasm
    - Connective_and_Soft_Tissue_Neoplasm
    - Endocrine_Neoplasm
    - Eye_Neoplasm
    - Gastrointestinal_Neoplasm
    - Head_and_Neck_Neoplasm
    - Hematopoietic_System_Neoplasm
    - Nervous_System_Neoplasm
    - Peritoneal_and_Retroperitoneal_Neoplasms
    - Reproductive_System_Neoplasm
    - Respiratory_Tract_Neoplasm
    - Skin_Neoplasm
      - Benign_Skin_Neoplasm
      - Cutaneous_Hematopoietic_and_Lymphoid_C
      - Dermal_Neoplasm
      - Epithelial_Skin_Neoplasm
      - Malignant_Skin_Neoplasm
      - Melanocytic_Skin_Neoplasm
      - Scrotal_Neoplasm
      - Skin_Appendage_Neoplasm

Melanoma

For Class: http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Malignant_Skin_Neo

Property	Value	Lang
DEFINITION	A primary or metastatic tumor involving the skin. Primary malignant skin tumors most often are carcinomas (either basal cell or squamous cell carcinomas that arise from cells in the epidermis) or melanomas that arise from pigment-containing skin melanocytes. Metastatic tumors to the skin include carcinomas and lymphomas.	
FULL_SYN	Melanoma and Non-Melanoma Skin Cancer	

NECESSARY & SUFFICIENT

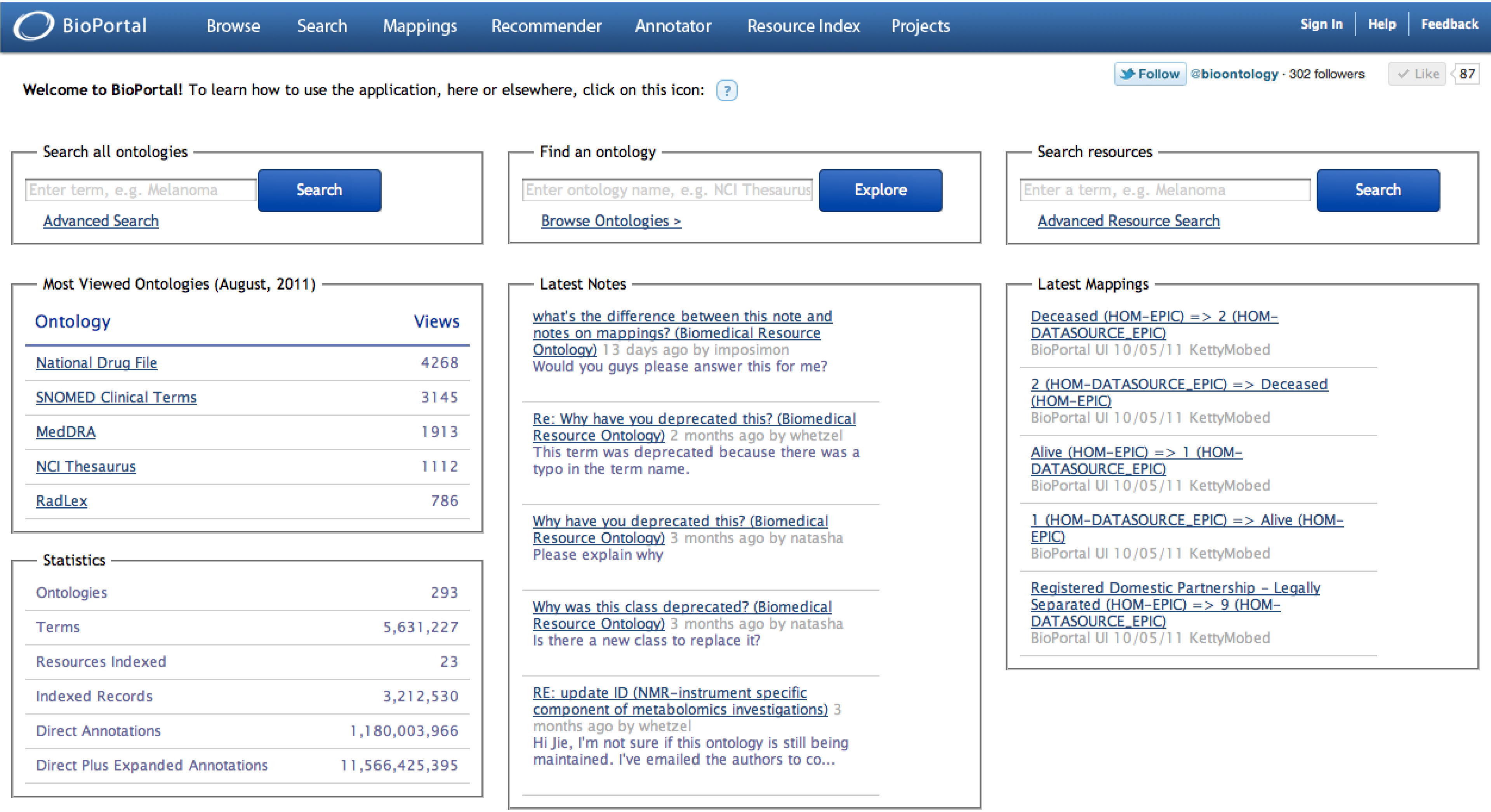
Skin_Neoplasm
Disease_Excludes_Normal_Cell_Origin only Transformed_Skin-Homing_T-Lymphocyte
Disease_Has_Abnormal_Cell only Malignant_Cell

NECESSARY

INHERITED

Disease_Has_Abnormal_Cell only Neoplastic_Cell [from Neoplasm]
Disease_Has_Associated_Anatomic_Site only Skin [from Skin_Neoplasm]
Disease_Has_Associated_Anatomic_Site only Integumentary_System [from Skin_Disorder]
Disease_Has_Finding only Cutaneous_Involvement [from Skin_Disorder]

BioPortal Ontology Repository



Welcome to BioPortal! To learn how to use the application, here or elsewhere, click on this icon: ?

Search all ontologies

Advanced Search

Find an ontology

Browse Ontologies >

Search resources

Advanced Resource Search

Most Viewed Ontologies (August, 2011)

Ontology	Views
National Drug File	4268
SNOMED Clinical Terms	3145
MedDRA	1913
NCI Thesaurus	1112
RadLex	786

Statistics

Ontologies	293
Terms	5,631,227
Resources Indexed	23
Indexed Records	3,212,530
Direct Annotations	1,180,003,966
Direct Plus Expanded Annotations	11,566,425,395

Latest Notes

what's the difference between this note and notes on mappings? (Biomedical Resource Ontology) 13 days ago by imposimon
Would you guys please answer this for me?

Re: Why have you deprecated this? (Biomedical Resource Ontology) 2 months ago by whetzel
This term was deprecated because there was a typo in the term name.

Why have you deprecated this? (Biomedical Resource Ontology) 3 months ago by natasha
Please explain why

Why was this class deprecated? (Biomedical Resource Ontology) 3 months ago by natasha
Is there a new class to replace it?

RE: update ID (NMR-instrument specific component of metabolomics investigations) 3 months ago by whetzel
Hi Jie, I'm not sure if this ontology is still being maintained. I've emailed the authors to co...

Latest Mappings

Deceased (HOM-EPIC) => 2 (HOM-DATASOURCE_EPIC)
BioPortal UI 10/05/11 KettyMobed

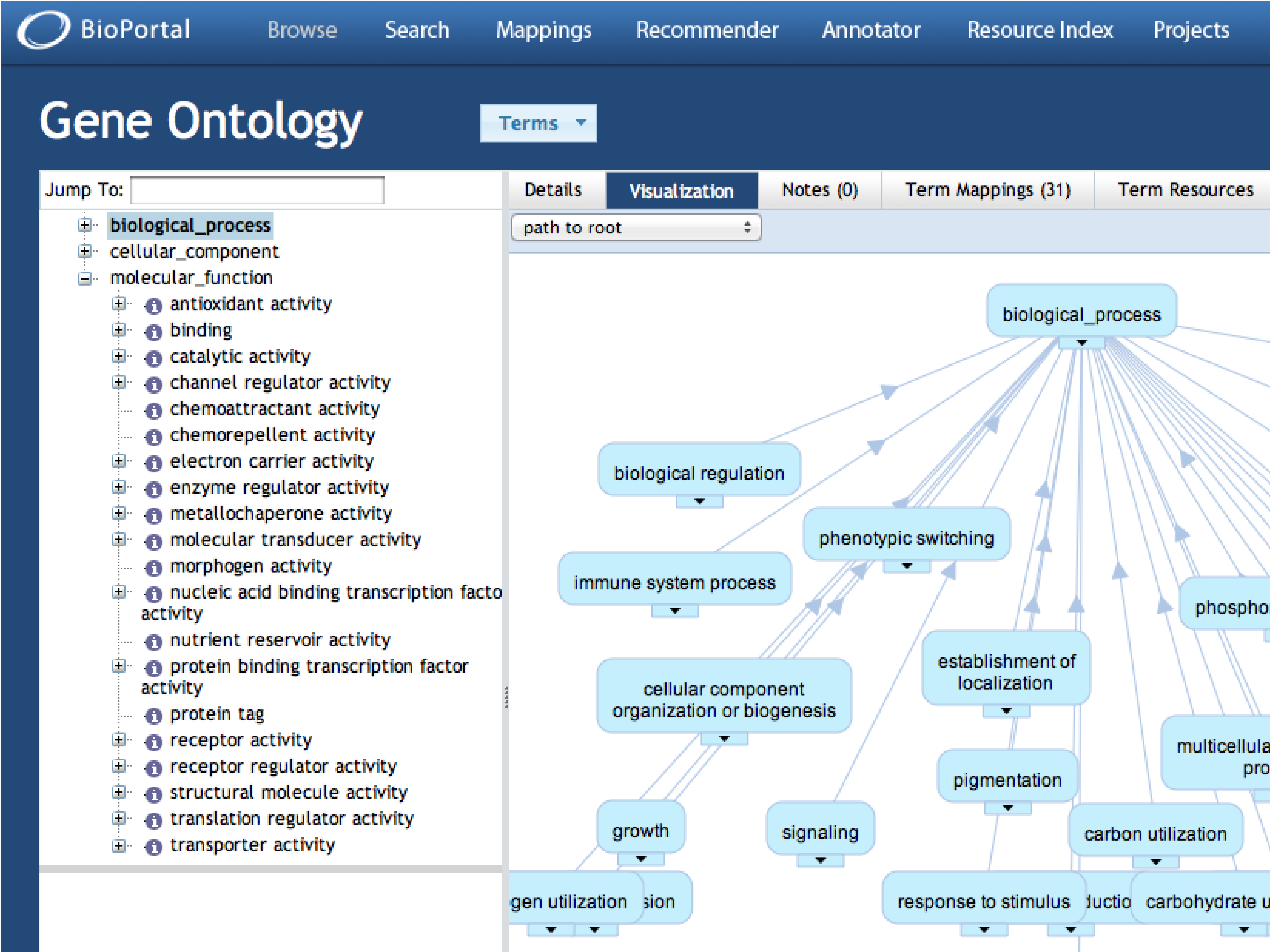
2 (HOM-DATASOURCE_EPIC) => Deceased (HOM-EPIC)
BioPortal UI 10/05/11 KettyMobed

Alive (HOM-EPIC) => 1 (HOM-DATASOURCE_EPIC)
BioPortal UI 10/05/11 KettyMobed

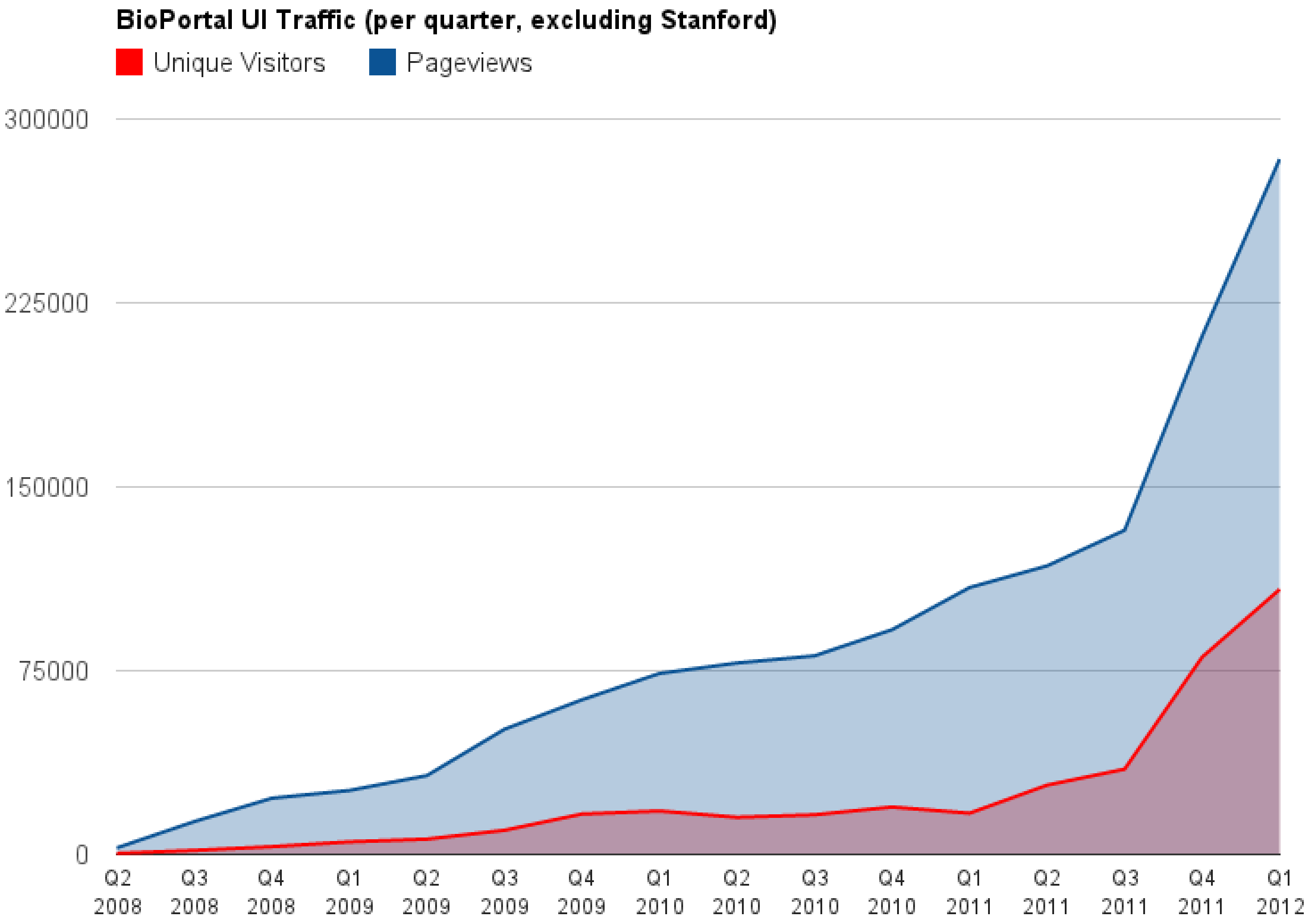
1 (HOM-DATASOURCE_EPIC) => Alive (HOM-EPIC)
BioPortal UI 10/05/11 KettyMobed

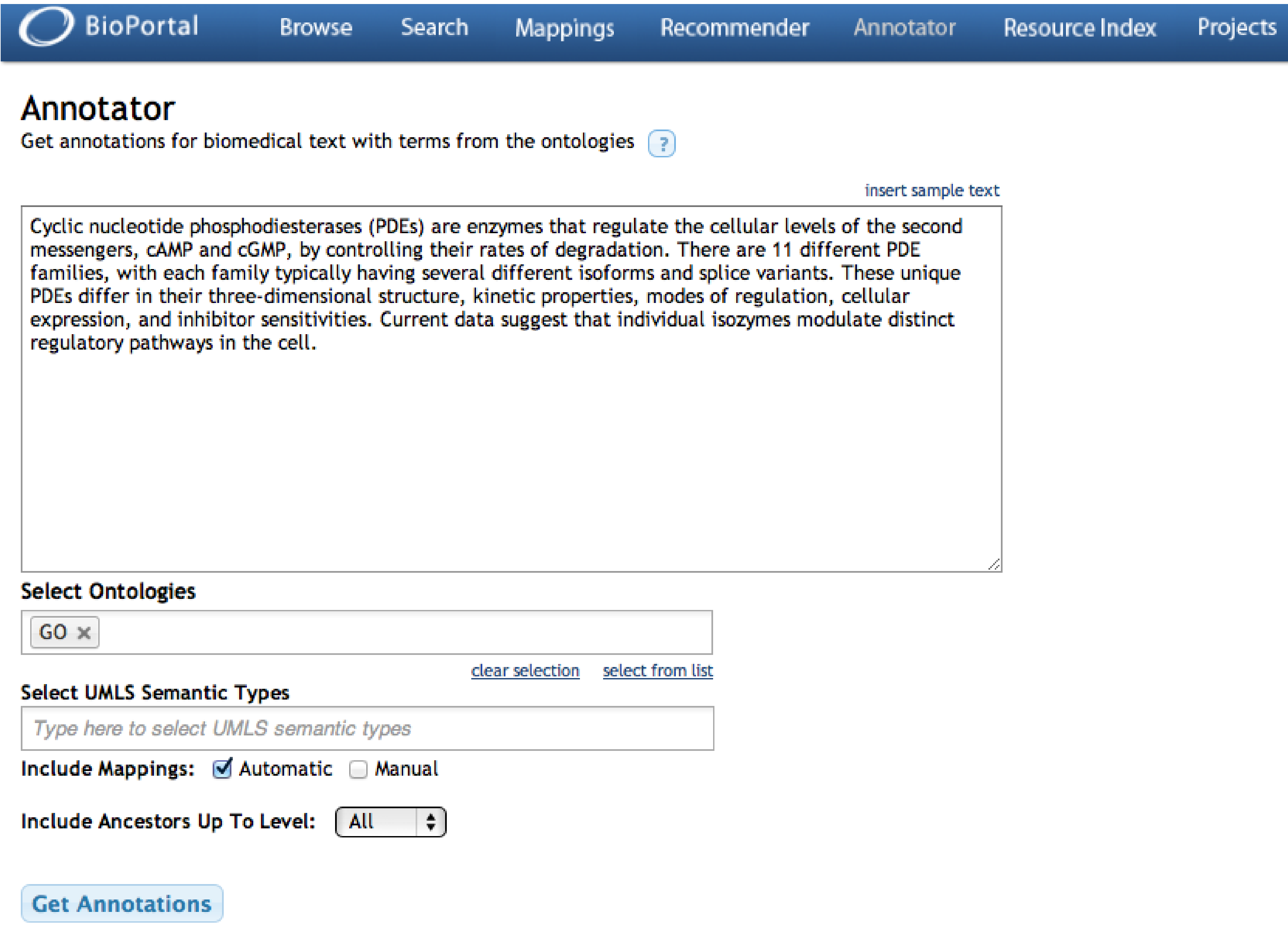
Registered Domestic Partnership - Legally Separated (HOM-EPIC) => 9 (HOM-DATASOURCE_EPIC)
BioPortal UI 10/05/11 KettyMobed

http://bioportal.bioontology.org


BioPortal
Browse
Search
Mappings
Recommender
Annotator
Resource Index
Projects
Gene Ontology
Terms
Jump To:
biological_process
cellular_component
molecular_function
antioxidant activity
binding
catalytic activity
channel regulator activity
chemoattractant activity
chemorepellent activity
electron carrier activity
enzyme regulator activity
metallochaperone activity
molecular transducer activity
morphogen activity
nucleic acid binding transcription facto activity
nutrient reservoir activity
protein binding transcription factor activity
protein tag
receptor activity
receptor regulator activity
structural molecule activity
translation regulator activity
transporter activity
Details
Visualization
Notes (0)
Term Mappings (31)
Term Resources
path to root
biological_process
biological regulation
phenotypic switching
immune system process
phospho
cellular component organization or biogenesis
establishment of localization
multicellula pro
pigmentation
growth
signaling
carbon utilization
gen utilization
sion
response to stimulus
ductio
carbohydrate u

BioPortal Traffic – Volume





Annotator

Get annotations for biomedical text with terms from the ontologies

Cyclic nucleotide phosphodiesterases (PDEs) are enzymes that regulate the cellular levels of the second messengers, cAMP and cGMP, by controlling their rates of degradation. There are 11 different PDE families, with each family typically having several different isoforms and splice variants. These unique PDEs differ in their three-dimensional structure, kinetic properties, modes of regulation, cellular expression, and inhibitor sensitivities. Current data suggest that individual isozymes modulate distinct regulatory pathways in the cell.

Select Ontologies

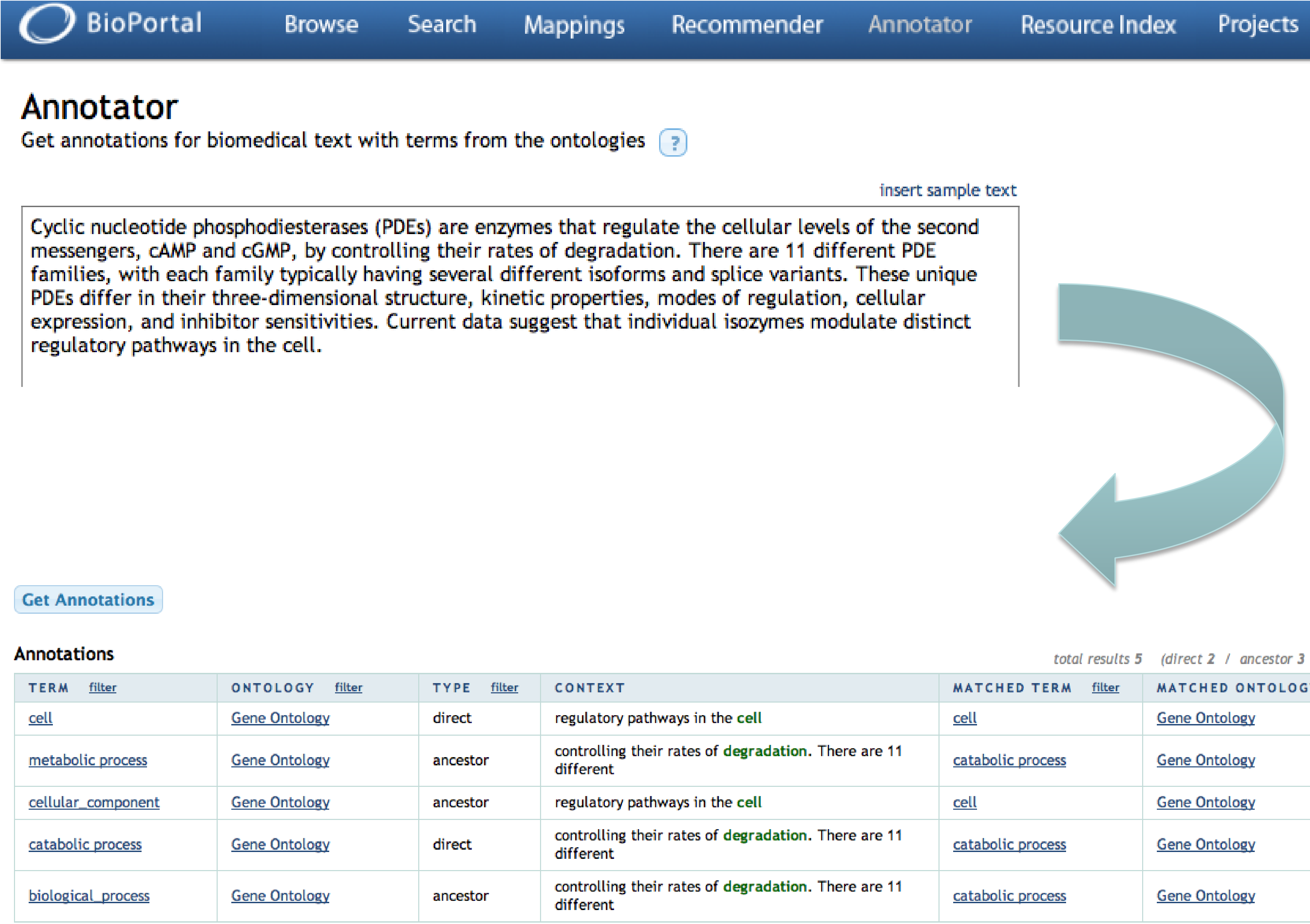
GO ×

Select UMLS Semantic Types

Type here to select UMLS semantic types

Include Mappings: ☑ Automatic ☐ Manual

Include Ancestors Up To Level: All



Annotator

Get annotations for biomedical text with terms from the ontologies

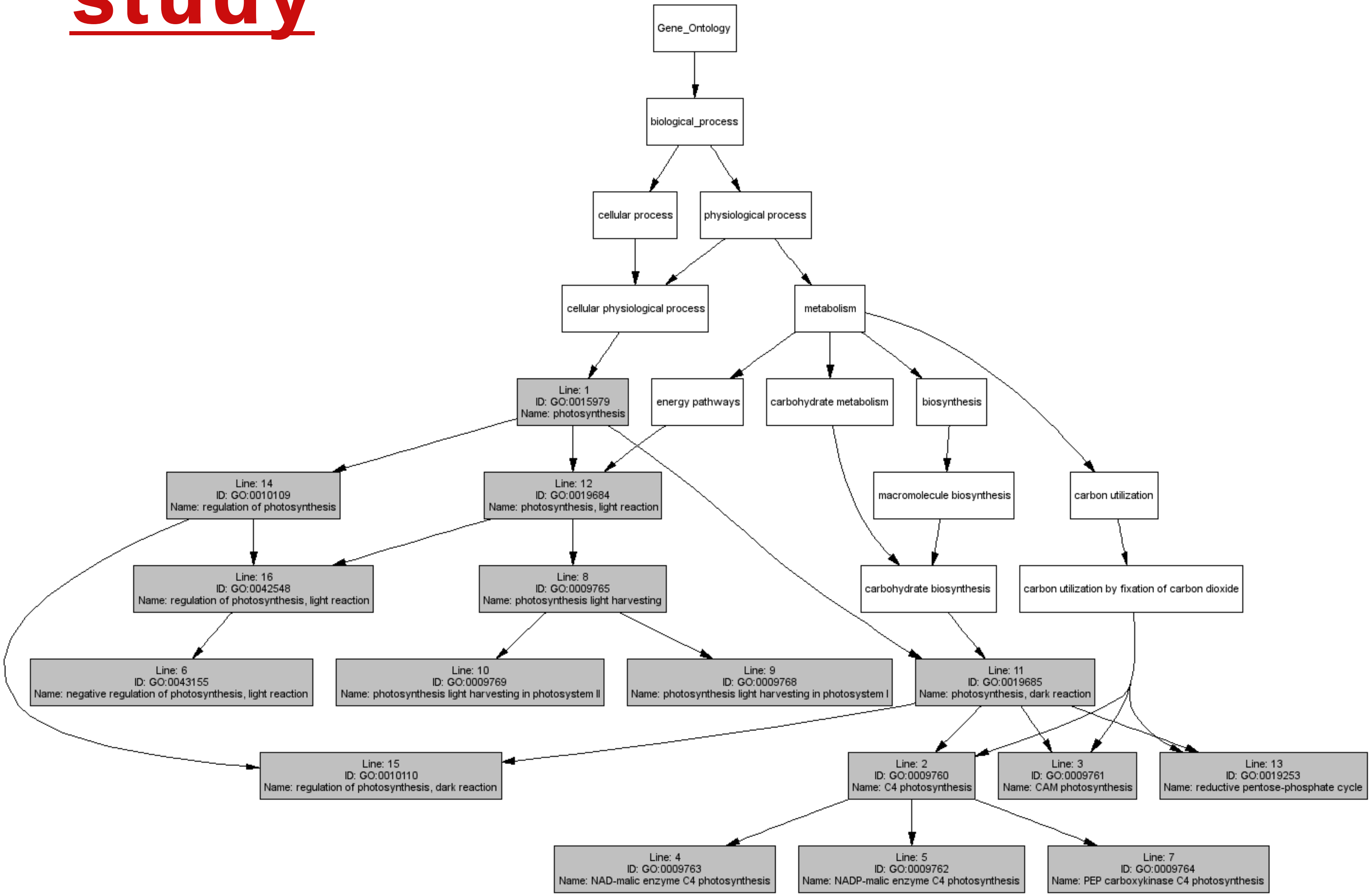
Cyclic nucleotide phosphodiesterases (PDEs) are enzymes that regulate the cellular levels of the second messengers, cAMP and cGMP, by controlling their rates of degradation. There are 11 different PDE families, with each family typically having several different isoforms and splice variants. These unique PDEs differ in their three-dimensional structure, kinetic properties, modes of regulation, cellular expression, and inhibitor sensitivities. Current data suggest that individual isozymes modulate distinct regulatory pathways in the cell.

Annotations

*total results **5** (direct **2** / ancestor **3***

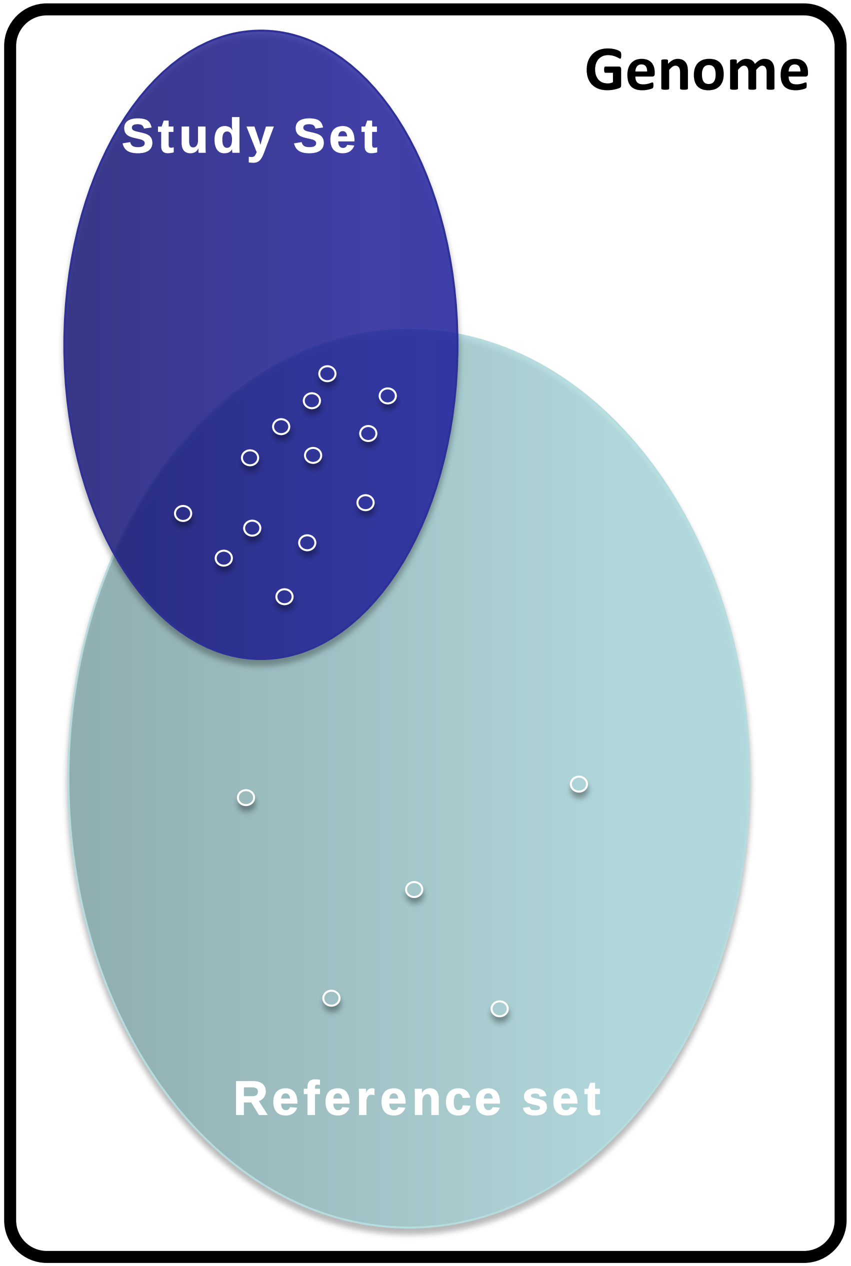
TERM filter	ONTOLOGY filter	TYPE filter	CONTEXT	MATCHED TERM filter	MATCHED ONTOLOG
cell	Gene Ontology	direct	regulatory pathways in the **cell**	cell	Gene Ontology
metabolic process	Gene Ontology	ancestor	controlling their rates of **degradation**. There are 11 different	catabolic process	Gene Ontology
cellular_component	Gene Ontology	ancestor	regulatory pathways in the **cell**	cell	Gene Ontology
catabolic process	Gene Ontology	direct	controlling their rates of **degradation**. There are 11 different	catabolic process	Gene Ontology
biological_process	Gene Ontology	ancestor	controlling their rates of **degradation**. There are 11 different	catabolic process	Gene Ontology

GO hierarchy can ascribe putative roles to genes in a study

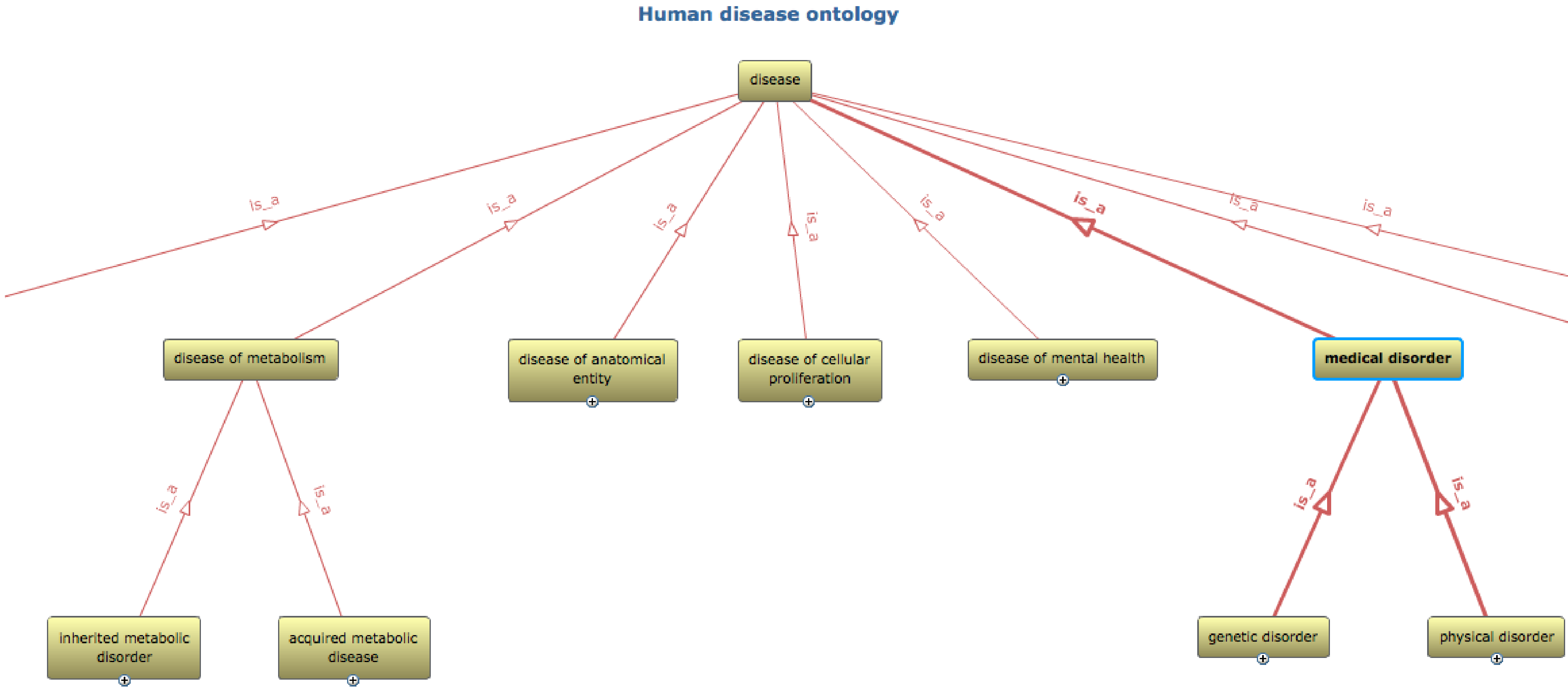


GO-based "enrichment analysis"

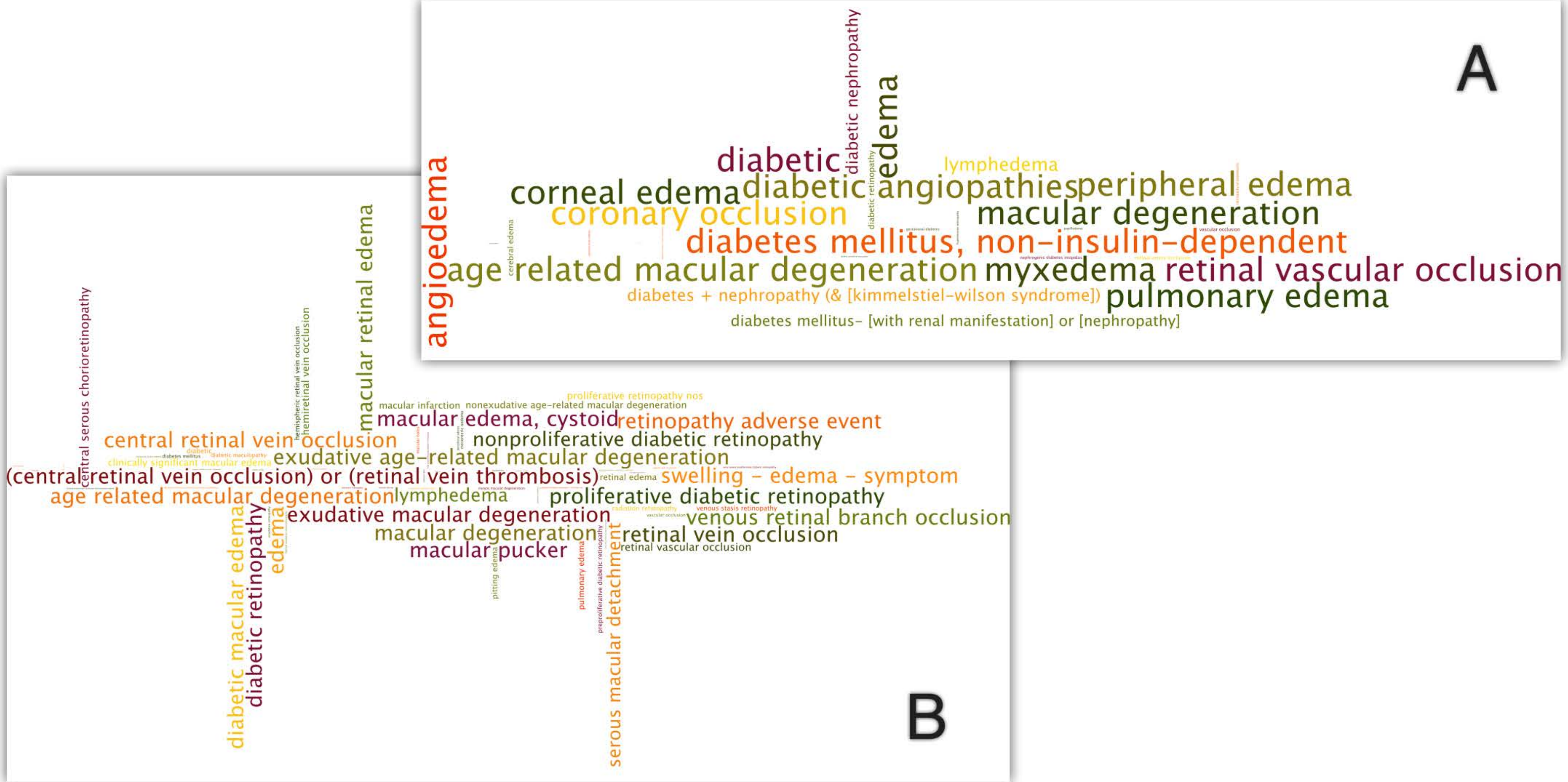
1. Get annotations for each gene in a set
2. Count the occurrence of each annotation term in the study set
3. Count the occurrence of that term in some reference set (whole genome?)
4. *p*-value for how surprising their overlap is



Suppose one had an ontology of diseases and some EHR data …



Shah et al. perform “enrichment analysis” on textual progress notes (in this case, for “Avastin”)



Thank you

For information contact Peter Lyster, Chair of the NIH Common Fund National Centers for Biomedical Computing Project Team

lysterp@mail.nih.gov
301.451.6446

or visit the consortium web site

http://www.ncbcs.org/