

BioPAX, Pathway Commons, Cytoscape and the Future of Scientific Publishing

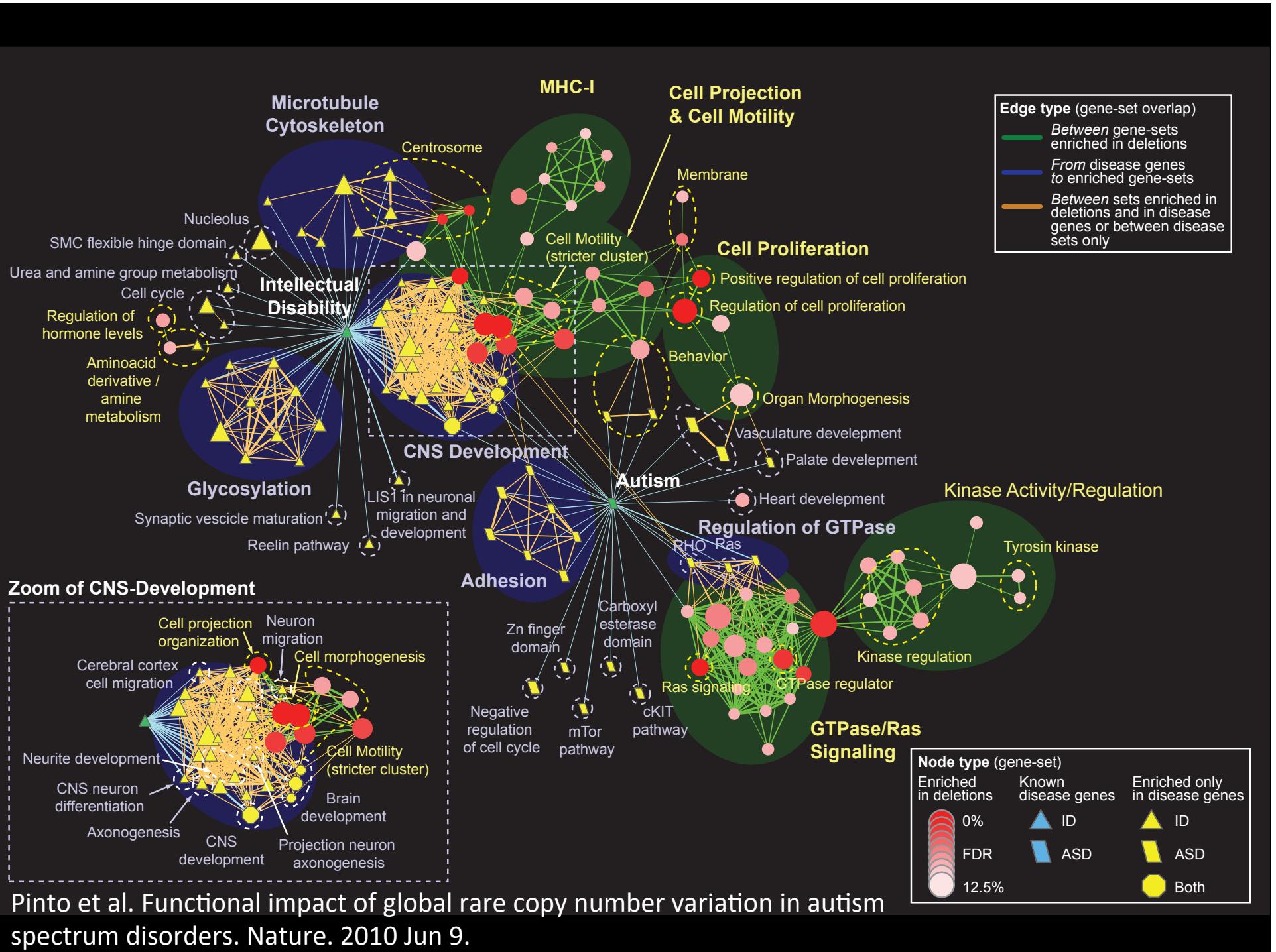
Gary Bader

May.21.2012 – HARMONY, Maastricht



Donnelly Centre
for Cellular + Biomolecular Research



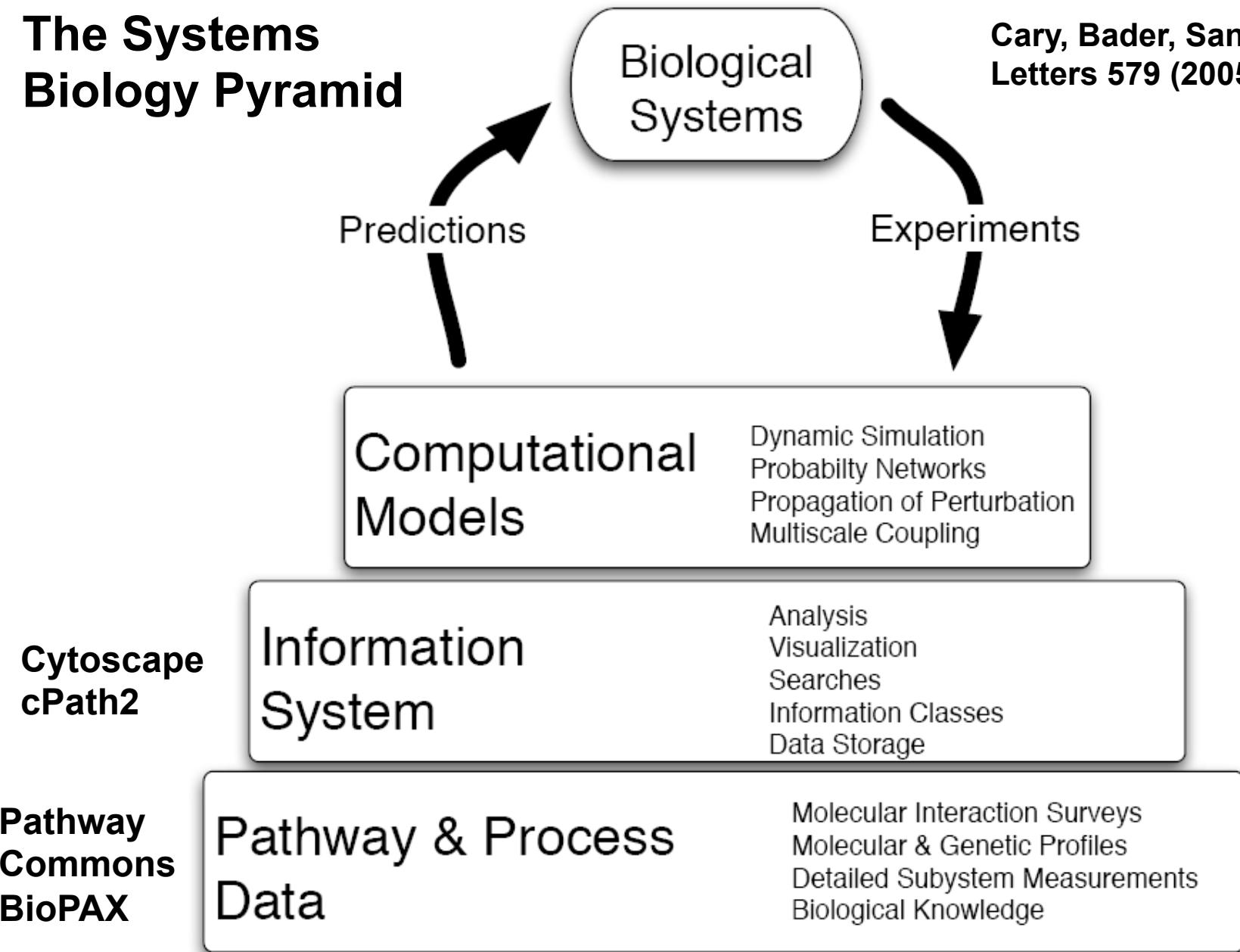


Correlation to Causation

- GWAS: find genetic markers correlated with disease – powerful approach, but:
 - genomics reduces statistical power (>multiple testing correction with >SNPs)
 - rare variants = more samples
- Associate pathways to increase power
 - Fewer pathways, organize many rare variants (damaging the system causes the disease)
- Use pathway knowledge to identify potential disease causes

The Systems Biology Pyramid

Cary, Bader, Sander, FEBS Letters 579 (2005) 1815-20



Pathguide

the pathway resource list

Navigation

- [Protein-Protein Interactions](#)
- [Metabolic Pathways](#)
- [Signaling Pathways](#)
- [Pathway Diagrams](#)
- [Transcription Factors / Gene Regulatory Networks](#)
- [Protein-Compound Interactions](#)
- [Genetic Interaction Networks](#)
- [Protein Sequence Focused](#)
- [Other](#)

Search

Organisms

All

Availability

All

Standards

All

[Reset](#) [Search](#)

Statistics

[Analyze Pathguide](#)

Contact

Comments, Questions,
Suggestions are
Always Welcome!

Complete Listing of All Pathguide Resources

Pathguide contains information about 222 biological pathway resources. Click on a link to go to the resource home page or 'Details' for a description page. Databases that are free and those supporting BioPAX, CellML, PSI-MI, or SBML standards are respectively indicated.

If you know of a pathway resource that is not listed here, or have other questions or comments, please [send us an e-mail](#).

[Home](#) [BioPAX](#) [cBio](#) [MSKCC](#)

>320 Pathway Databases!

Get the Stats

Detailed Pathguide resource statistics now available

Pathguide Published

Please cite the [Pathguide](#)

Protein-Protein Interactions

Database Name (Order: alphabetically | by web popularity 

	Full Record	Availability	Standards
3DID - 3D interacting domains	Details	 Free	
ABCdb - Archaea and Bacteria ABC transporter database	Details	 Free	
AfCS - Alliance for Cellular Signaling Molecule Pages Database	Details	 Free	
AllFuse - Functional Associations of Proteins in Complete Genomes	Details	 Free	
ASEdb - Alanine Scanning Energetics Database	Details	 Free	
ASPD - Artificial Selected Proteins/Peptides Database	Details	 ?	
BID - Binding Interface Database	Details	 Free	
BIND - Biomolecular Interaction Network Database	Details	 Free	
BindingDB - The Binding Database	Details	 Free	
BioGRID - General Repository for Interaction Datasets	Details	 Free	
BRITE - Biomolecular Relations in Information Transmission and Expression	Details	 Free	
CA1Neuron - Pathways of the hippocampal CA1 neuron	Details	 Free	
Cancer Cell Map - The Cancer Cell Map	Details	 Free	
CSP - Cytokine Signaling Pathway Database	Details	 Free	
CTDB - Calmodulin Target Database	Details	 Free	
DDIB - Database of Domain Interactions and Bindings	Details	 Free	
DIP - Database of Interacting Proteins	Details	 Free	
Doodle - Database of oligomerization	Details	 Free	
DopaNet - DopaNet	Details	 Free	
DRC - Database of Ribosomal RNA Processing	Details	 Free	
DSM - Dynamic Signaling Maps	Details	 Free	
FIMM - Functional Molecular Interaction Maps	Details	 Free	
FusionDB - Prokaryote Gene Fusion Database	Details	 Free	

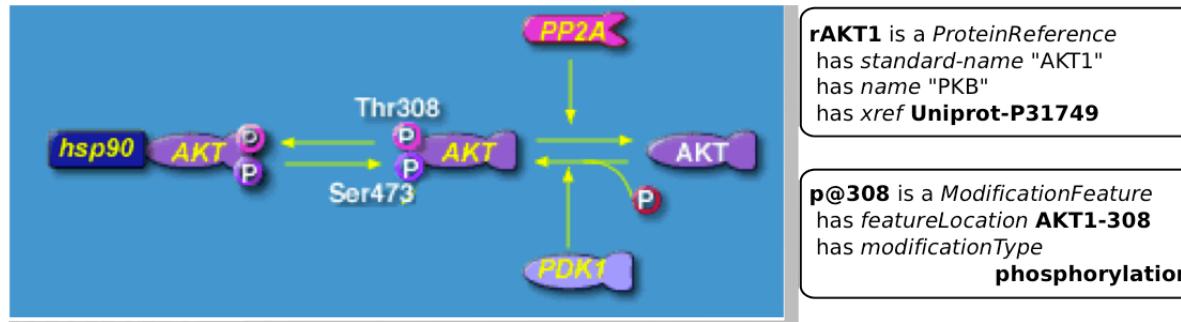
- Varied formats, representation, coverage
- Pathway data extremely difficult to combine and use

BioPAX Pathway Language

- Represent:
 - Metabolic pathways
 - Signaling pathways
 - Protein-protein, molecular interactions
 - Gene regulatory pathways
 - Genetic interactions
- Community effort: pathway databases distribute pathway information in standard format

www.biopax.org

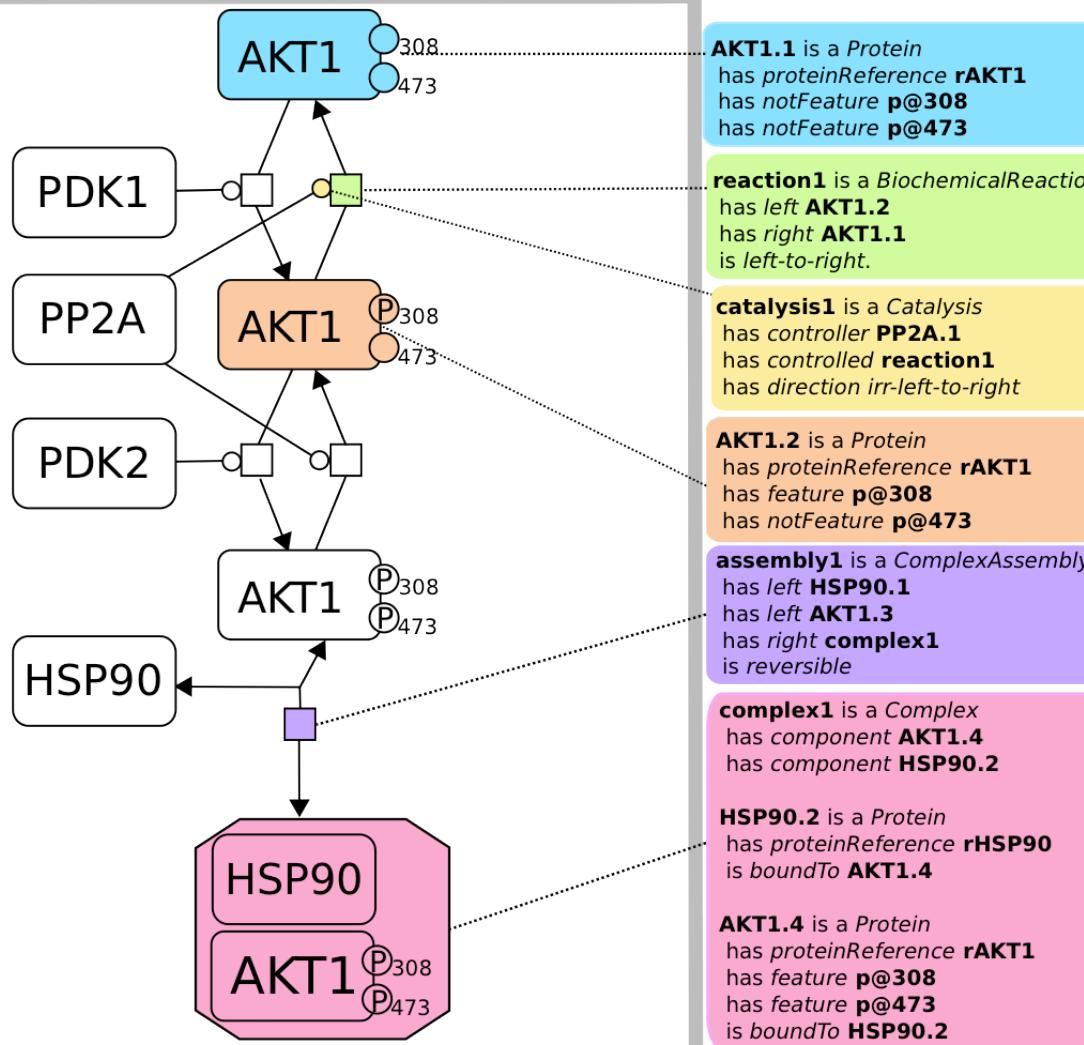
BioCarta



rAKT1 is a *ProteinReference*
has *standard-name* "AKT1"
has *name* "PKB"
has *xref* [Uniprot-P31749](#)

p@308 is a *ModificationFeature*
has *featureLocation* **AKT1-308**
has *modificationType*
phosphorylation

SBGN.org



AKT1.1 is a *Protein*
has *proteinReference* **rAKT1**
has *notFeature* **p@308**
has *notFeature* **p@473**

reaction1 is a *BiochemicalReaction*
has *left* **AKT1.2**
has *right* **AKT1.1**
is *left-to-right*.

catalysis1 is a *Catalysis*
has *controller* **PP2A.1**
has *controlled* **reaction1**
has *direction* *irr-left-to-right*

AKT1.2 is a *Protein*
has *proteinReference* **rAKT1**
has *feature* **p@308**
has *notFeature* **p@473**

assembly1 is a *ComplexAssembly*
has *left* **HSP90.1**
has *left* **AKT1.3**
has *right* **complex1**
is *reversible*

complex1 is a *Complex*
has *component* **AKT1.4**
has *component* **HSP90.2**

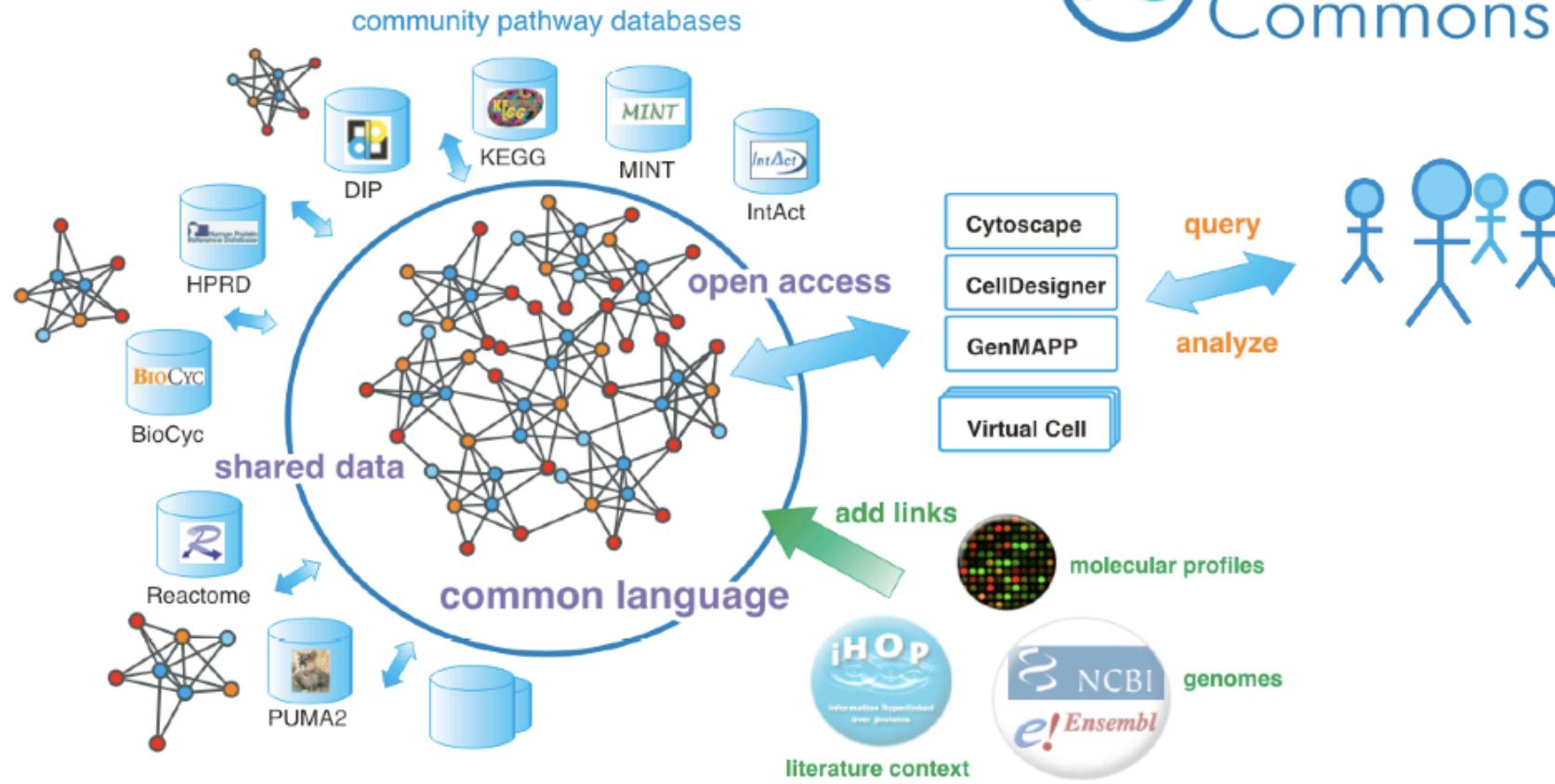
HSP90.2 is a *Protein*
has *proteinReference* **rHSP90**
is *boundTo* **AKT1.4**

AKT1.4 is a *Protein*
has *proteinReference* **rAKT1**
has *feature* **p@308**
has *feature* **p@473**
is *boundTo* **HSP90.2**

BioPAX

Aim: Convenient Access to Pathway Information

<http://pathwaycommons.org>



Facilitate creation and communication of pathway data
Aggregate pathway data in the public domain
Provide easy access for pathway analysis

Long term: Converge
to integrated cell map

Pathway Commons: cPath²

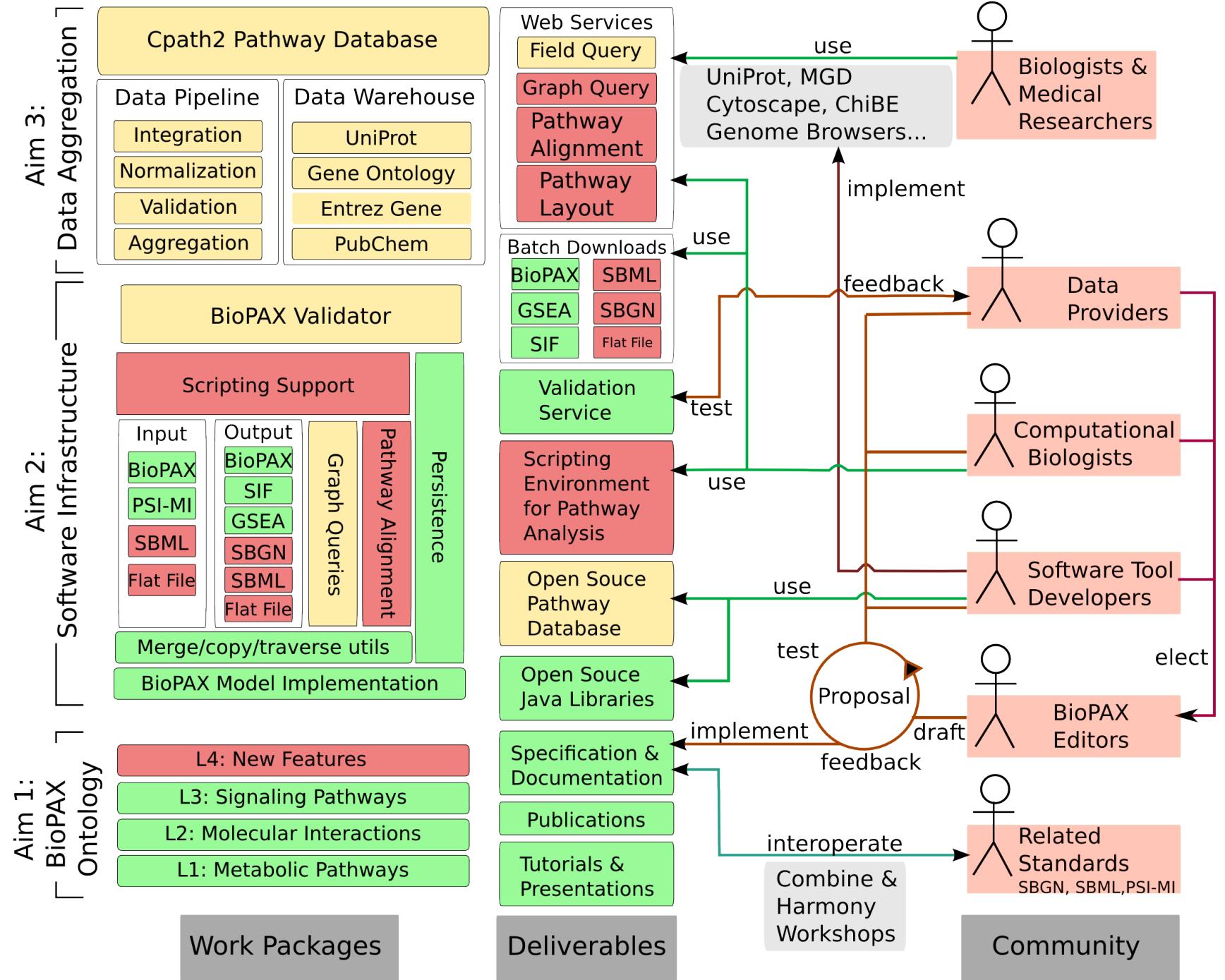
- <http://www.pathwaycommons.org/pc2-demo/>

Web Service API:

You can programmatically access the data within Pathway Commons using the Pathway Commons Web Service Application Programming Interface (API). This page provides a reference guide to help you get started.

1. [Command: SEARCH](#)
2. [Command: GET](#)
3. [Command: GRAPH](#)
4. [Command: TRAVERSE](#)
5. [Command: TOP PATHWAYS](#)
6. [Command: HELP](#)

**Emek Demir, Igor Rodchenkov, Chris Sander Ozgun Babur, Arman Aksoy,
Onur Sumer, Ethan Cerami, Ben Gross**



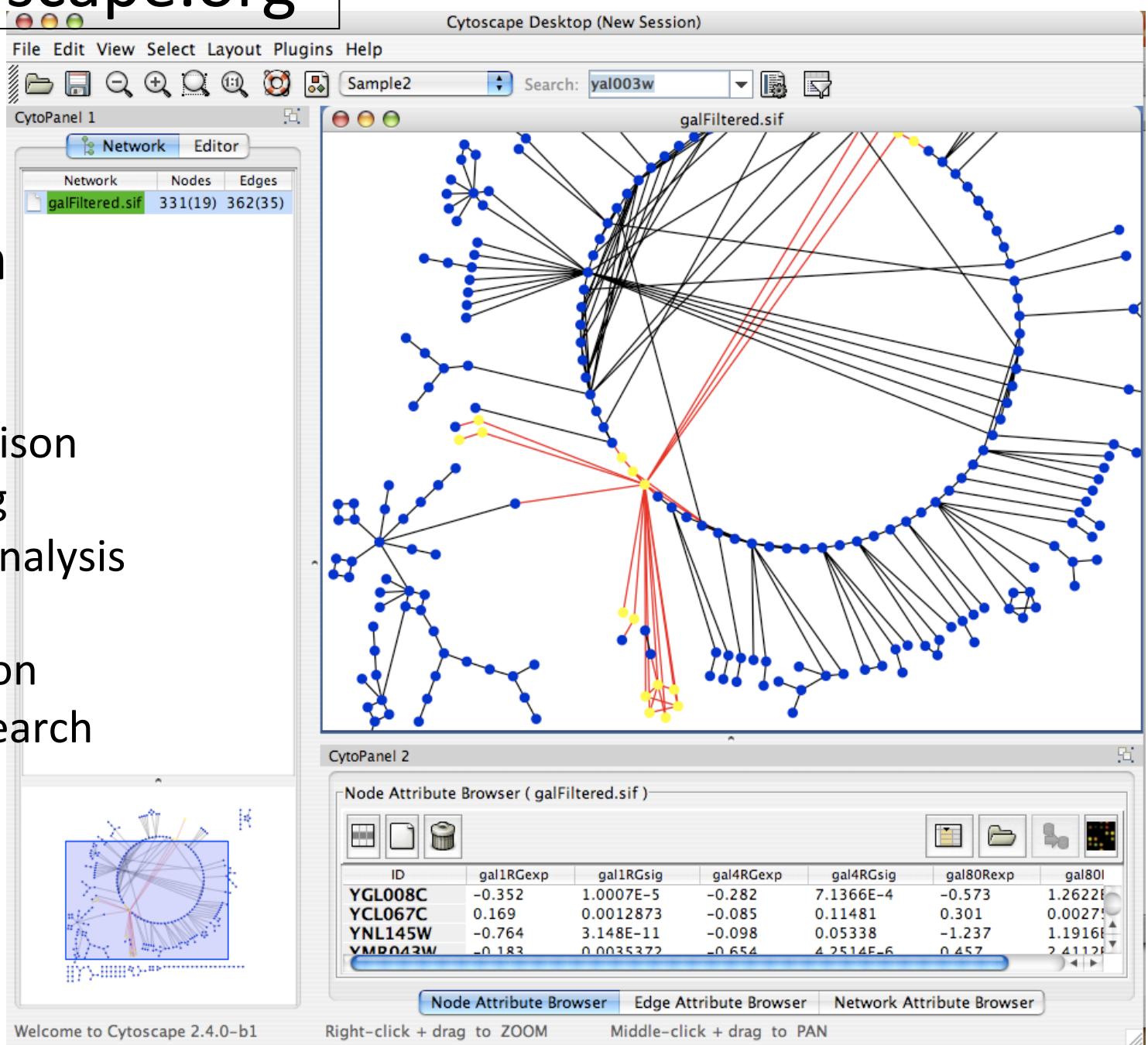
Aim 4: Community Organization and Dissemination

<http://cytoscape.org>

Network visualization and analysis

Pathway comparison
Literature mining
Gene Ontology analysis
Active modules
Complex detection
Network motif search

UCSD, ISB, Agilent,
MSKCC, Pasteur, UCSF,
Unilever, UToronto, U
Texas



Cytoscape 3

- Complete re-architecture: OSGi – everything is an app
- Enables future features:
 - More stable and powerful APIs
 - Scripting, macros, recordable history, better undo/redo
 - Command line mode, good for use on compute clusters
 - Interactive control from other scripting languages e.g. R
- Fixing bugs and porting plugins
- 3.0 developer beta now available
 - Mirror functionality in 2.8
 - Encourage plugin porting
 - http://www.cytoscape.org/documentation_cy3_dev.html

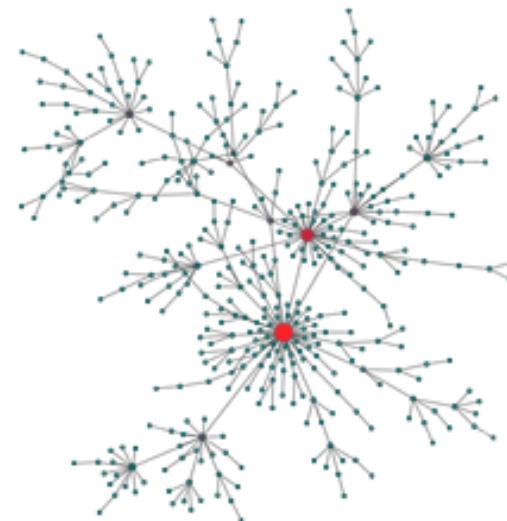


NATIONAL RESOURCE FOR NETWORK BIOLOGY

Mission Statement

The aim of the National Resource for Network Biology (NRNB) is to provide a freely available, open-source suite of software technology that broadly enables network-based visualization, analysis, and biomedical discovery for NIH-funded researchers. This software is enabling researchers to assemble large-scale biological data into models of networks and pathways and to use these networks to better understand how biological systems operate under normal conditions and how they fail in disease.

The National Resource for Network Biology is an NIH National Center for Research Resources (**NCRR**) Biomedical Technology Research Center (**BTRC**), organized around the following key components: Technology Research and Development, Driving Biomedical Projects, Outreach, Training and Dissemination of Tools. NRNB is funded by RR031228.



What is Network Biology?

Collaborate with NRNB

[View ongoing collaborations](#)

Annual Report



Cytoscape
Web

Feature S

This is a sep
Because this

<http://cytoscapeweb.cytoscape.org/>

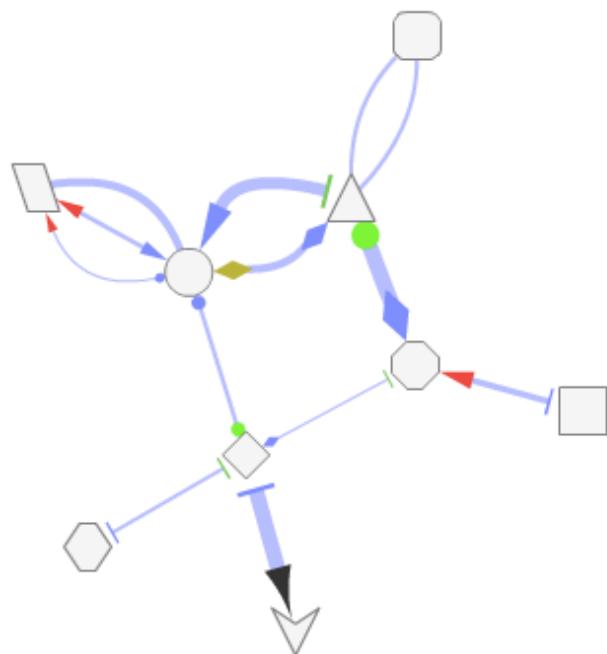
This showcase is complex, you may experience issues, such as slowdowns, on older or less efficient browsers.

Save file

Open file

Style ▾

Layout ▾



Examples

Visual style

Filter

Properties

Shapes example



A graph that contains all possible shapes for nodes and arrows

Petersen example



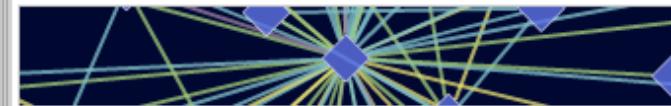
The Petersen graph

Disconnected example



A graph that contains several, disconnected components

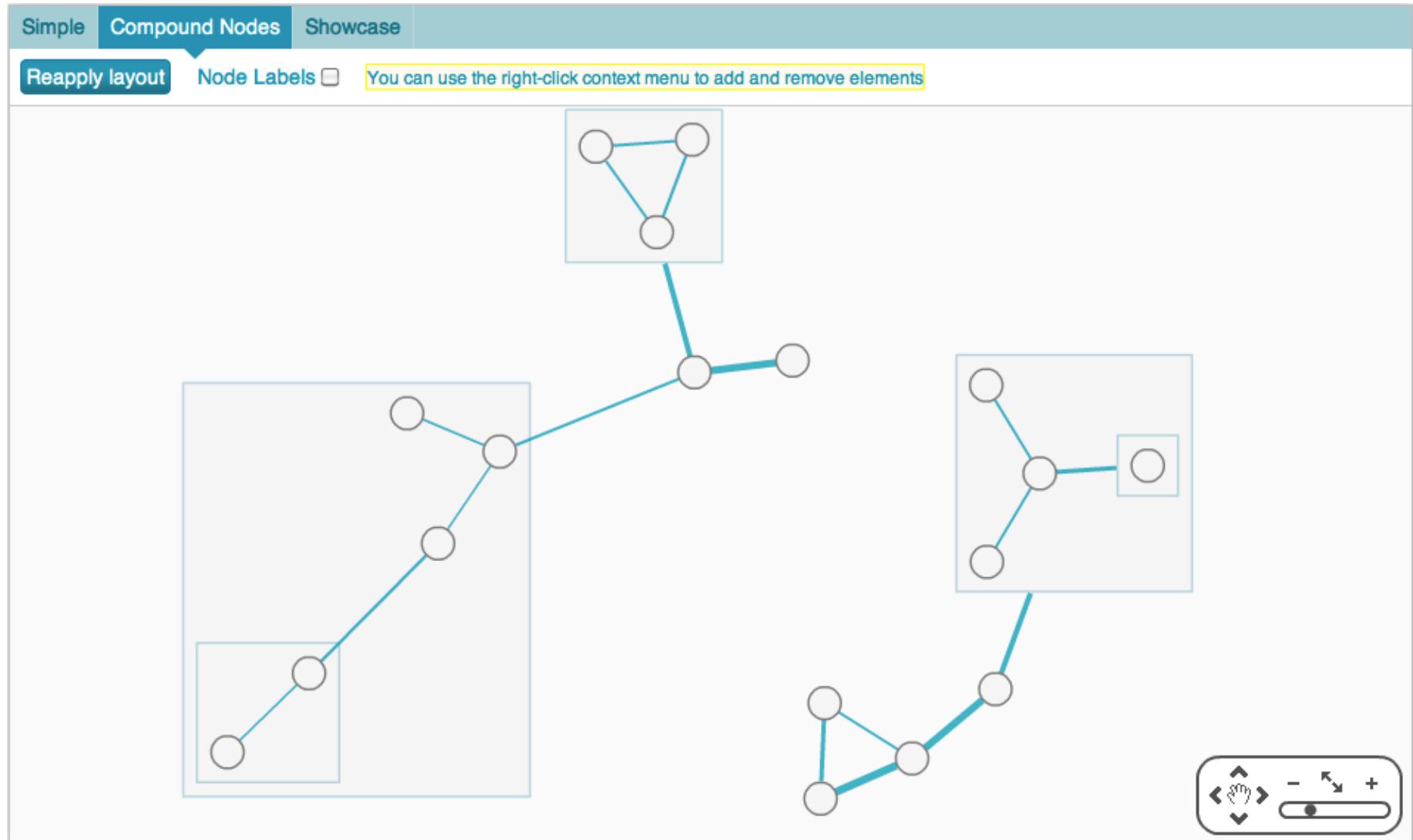
Genetics example



A modified graph from GeneMANIA with different visual styles

Onur Sumer
Ugur Dogrusoz
Bilkent, Ankara

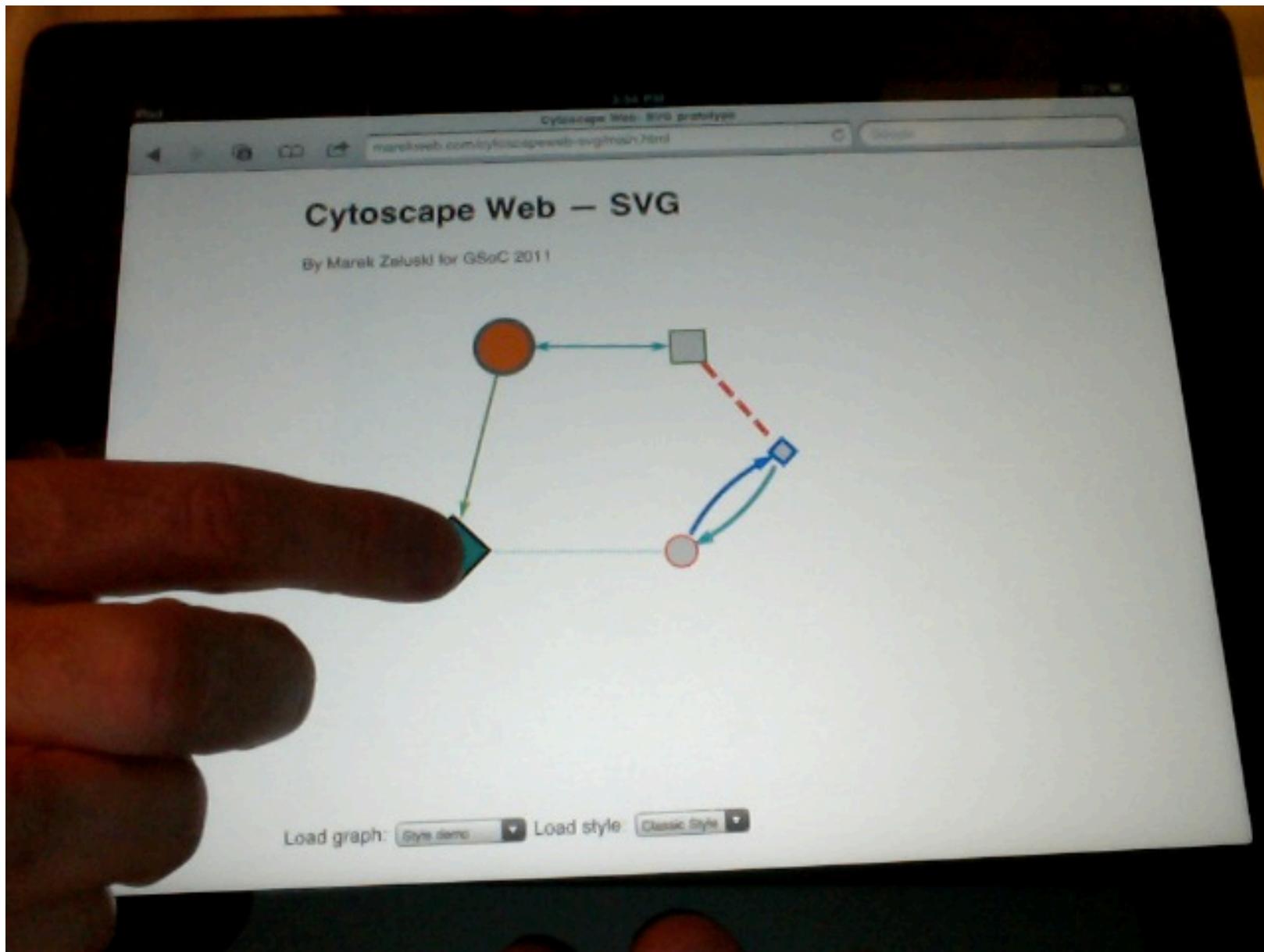
Compound Nodes



<http://cytoscapeweb.cytoscape.org/demos/compound>

<http://cytoscape.github.com/cytoscapeweb/>

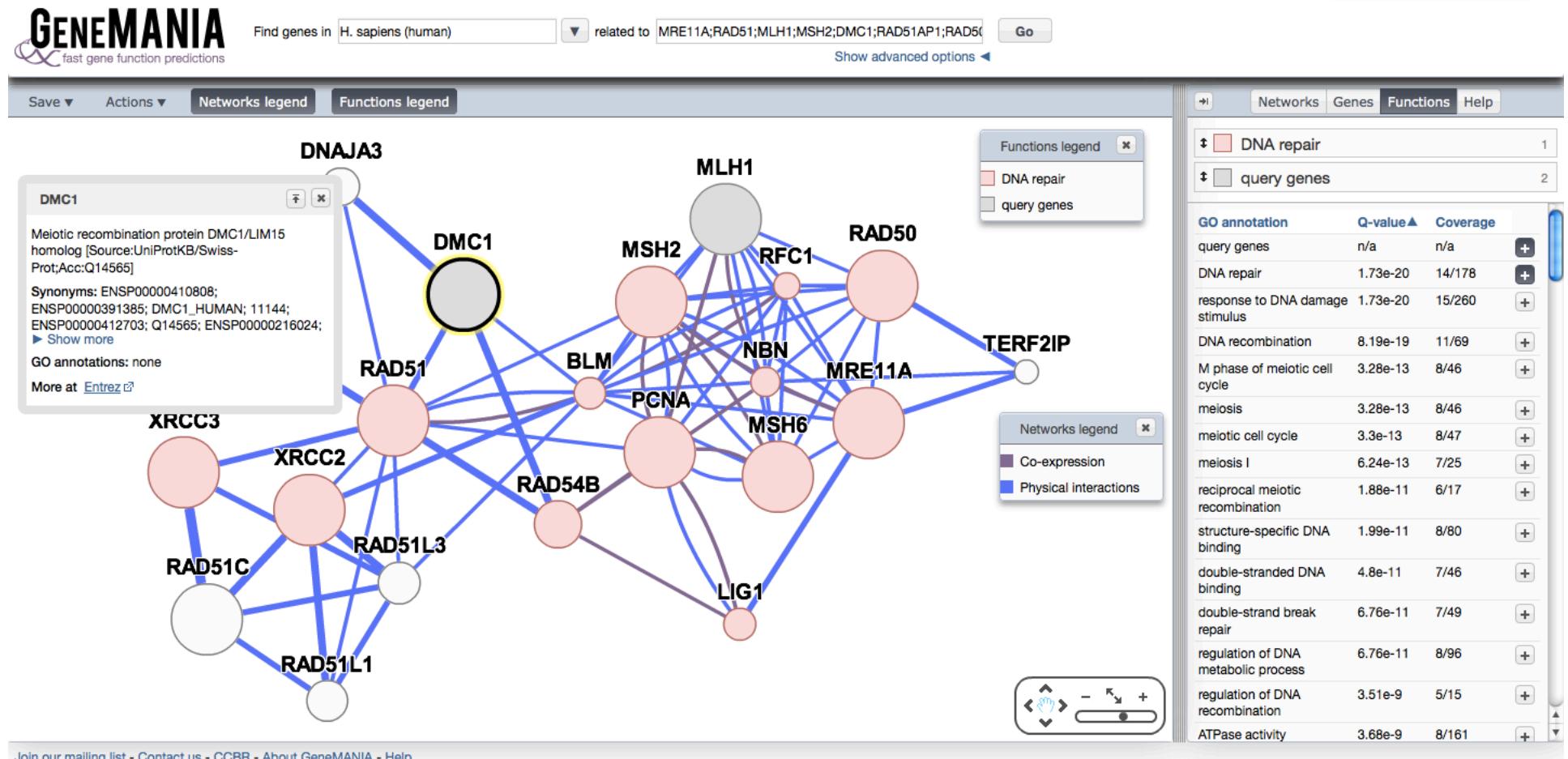
Cytoscape Web 2: HTML5 SVG – iPad!



<http://www.genemania.org>

Mostafavi S et al. Genome Biol. 2008;9 Suppl 1:S4

Warde-Farley D et al. Nucleic Acids Res. 2010 Jul; 8:W214-20.



[Join our mailing list](#) - [Contact us](#) - [CCBR](#) - [About GeneMANIA](#) - [Help](#)

- Guilt-by-association principle
- Biological networks are combined intelligently to optimize prediction accuracy
- Algorithm is more fast and accurate than its peers

Gene Function Prediction

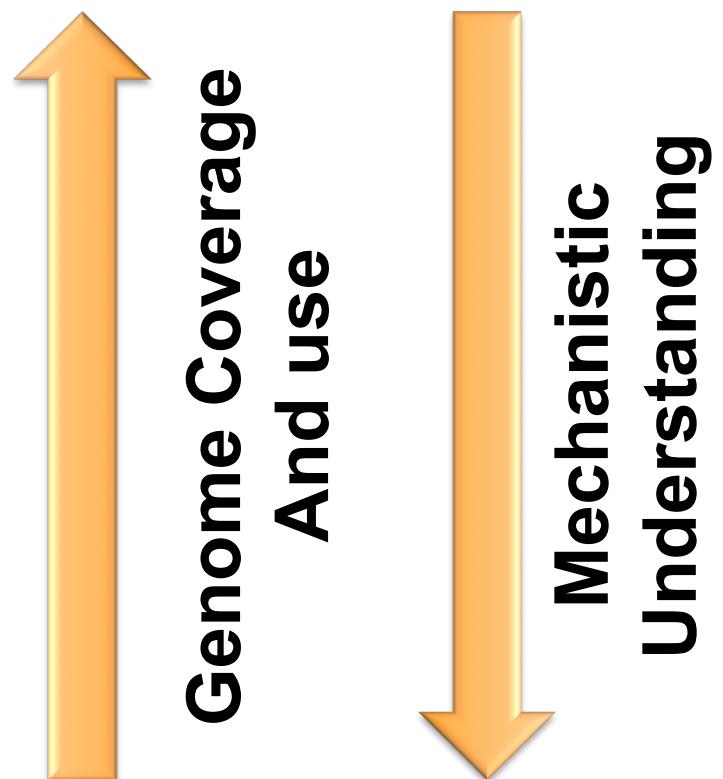
Quaid Morris (Donnelly), Sara Mostafavi
Rashad Badrawi, Ovi Comes, Sylva Donaldson,
Max Franz, Christian Lopes, Farzana Kazi,
Jason Montojo, Harold Rodriguez, Khalid Zuberi

The Factoid Project

- Publishing in science
 - Highly inefficient
 - Outdated technology, difficult to search and compute
- <http://www.elseviergrandchallenge.com/>
 - Winner: <http://reflect.ws/>
- Pathway and network information database curation
 - Highly inefficient
- The factoid project

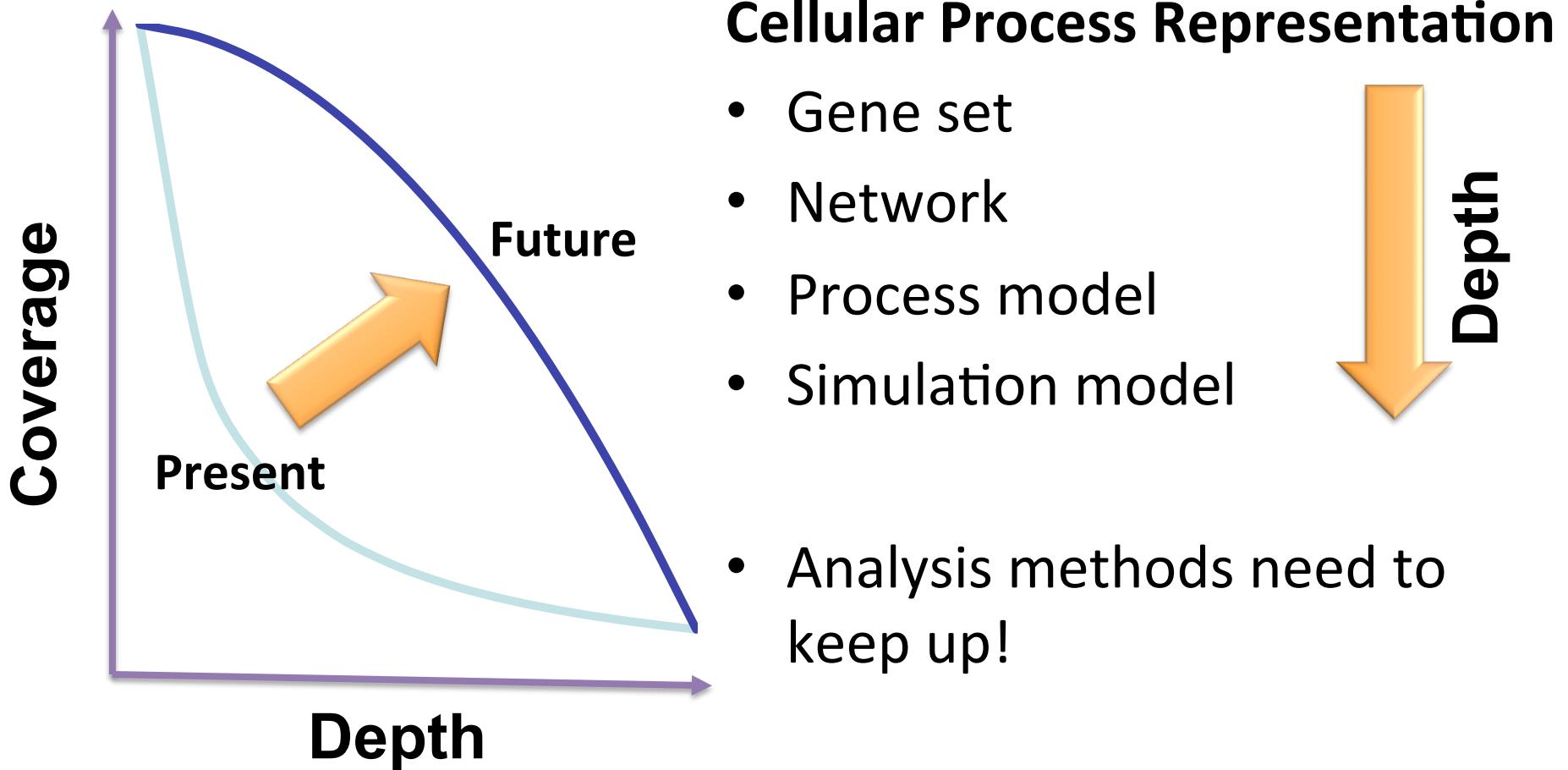
Pathway and Network Analysis

1. Gene set: pathway enrichment analysis
2. Network: network regions (modules), regulation
3. Process model: classical pathways
4. Simulation model: detailed models

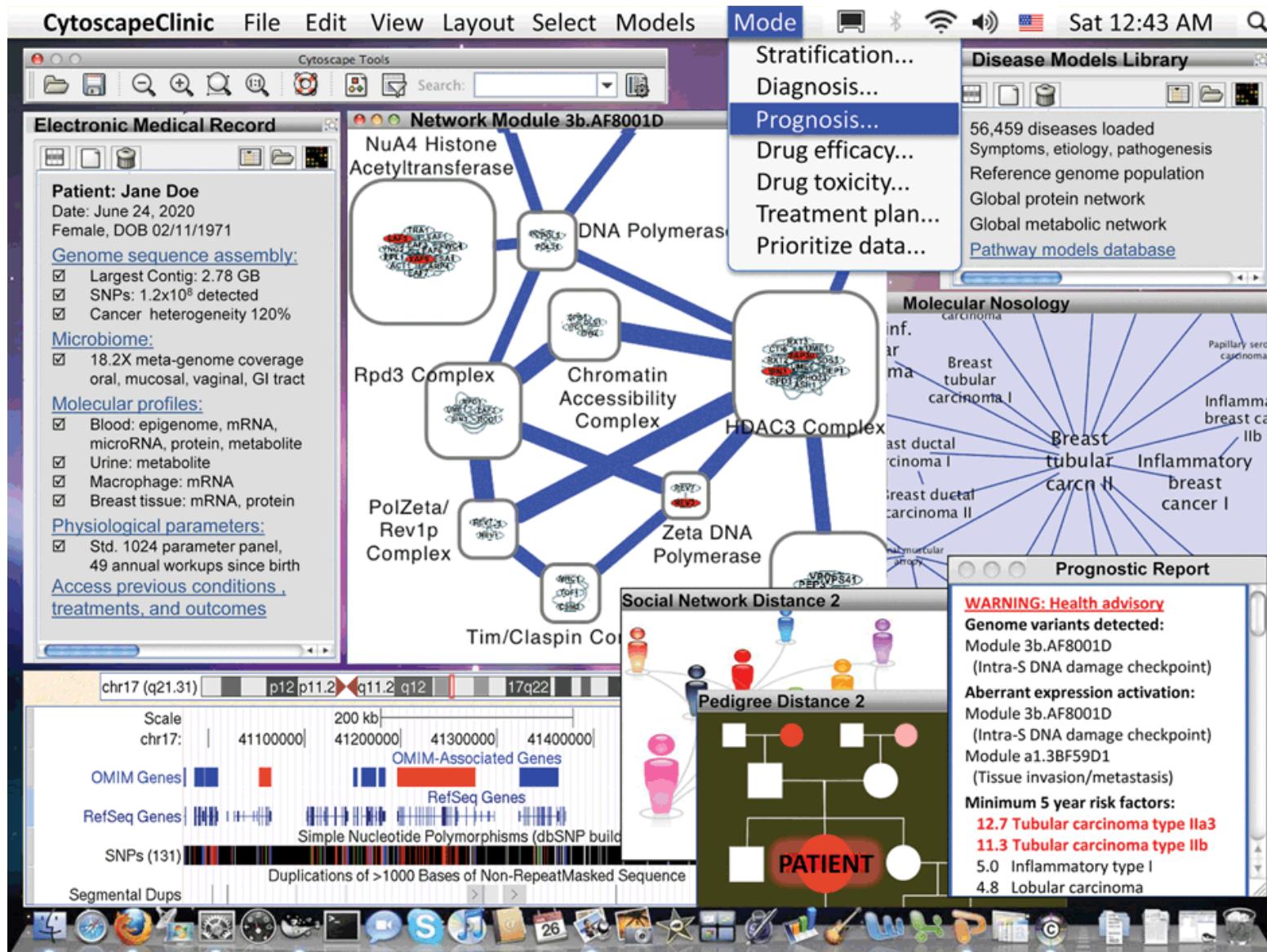


Increase Coverage and Depth

Data and analysis methods



Are we prepared for the future doctor visit?



Point: Are we prepared for the future doctor visit? Friend SH, Ideker T. Nat Biotechnol. 2011 Mar;29

ICSB – Systems Biology, Toronto

ICSB

THE 13th INTERNATIONAL
CONFERENCE ON
SYSTEMS BIOLOGY



General Information

Plenary Speakers

Fees and Registration

Scientific Topics

Exhibitor and Sponsor
Information

Abstract Submission

Accommodation

Social and Tour
Programme

Schedule

Travel Information

Contact Information

AUGUST . 19–23 . 2012

COMBINE
August
15-19, 2012
UToronto



Acknowledgements

Bader Lab

Domain Interaction Team

Chris Tan

Shirley Hui

Shobhit Jain

Brian Law

Jüri Reimand

Former:

David Gfeller

Xiaojian Shao

Genetic Intx, Pathways:

Anastasia Baryshnikova

Iain Wallace

Magali Michaut

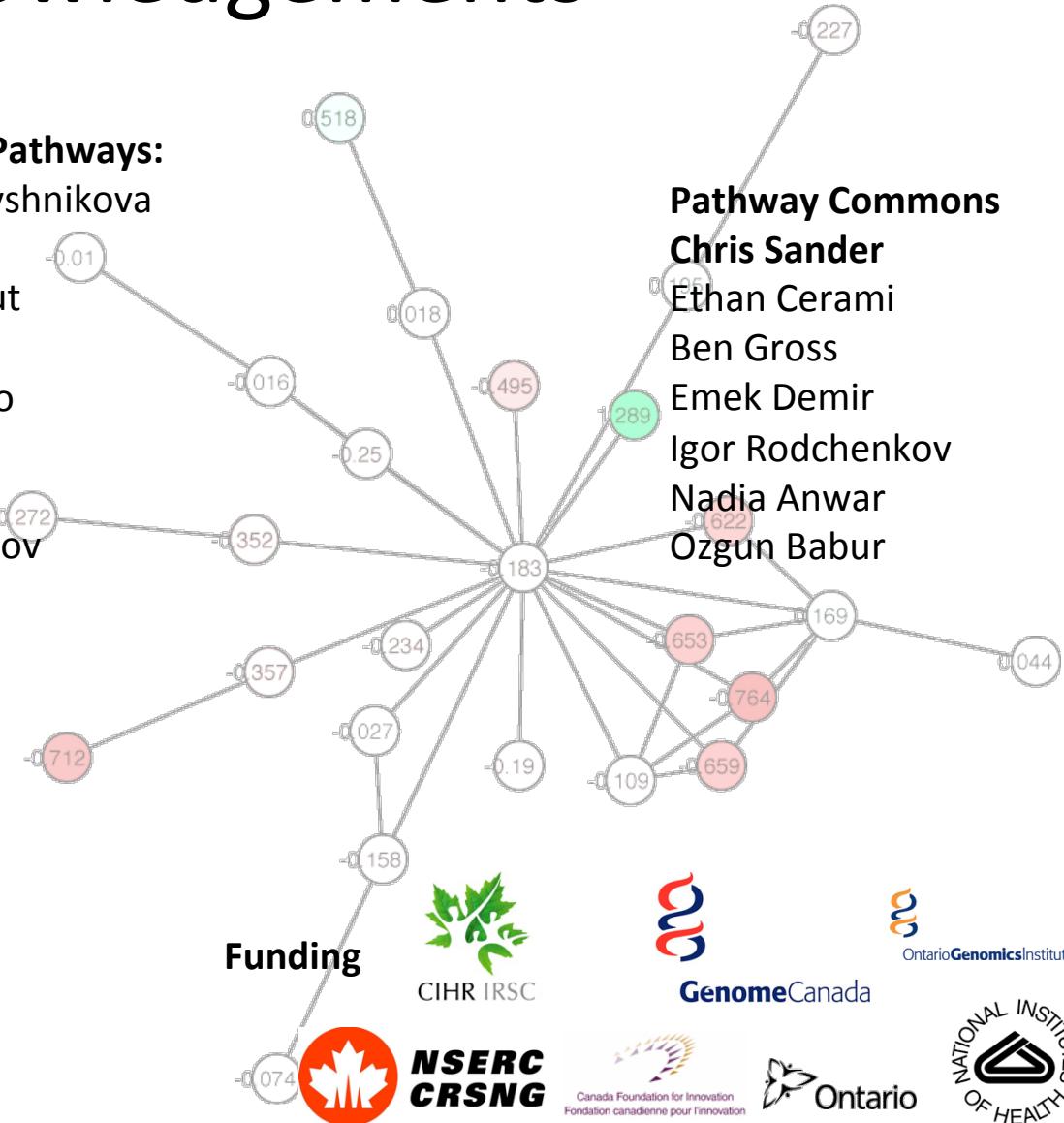
Ron Ammar

Daniele Merico

Ruth Isserlin

Vuk Pavlovic

Igor Rodchenkov



www.GeneMANIA.org
Quaid Morris (Donnelly)
Rashad Badrawi, Ovi
Comes, Sylva
Donaldson,
Christian Lopes,
Farzana Kazi, Jason
Montojo, Sara Mostafavi,
Harold Rodriguez,
Khalid Zuberi

<http://baderlab.org>