

Multiple Perturbation Mapping of Biological Systems

Gary Bader

Donnelly Centre for Cellular and Biomolecular Research

University of Toronto

Centennial Bioinformatics, Mar.2 2009

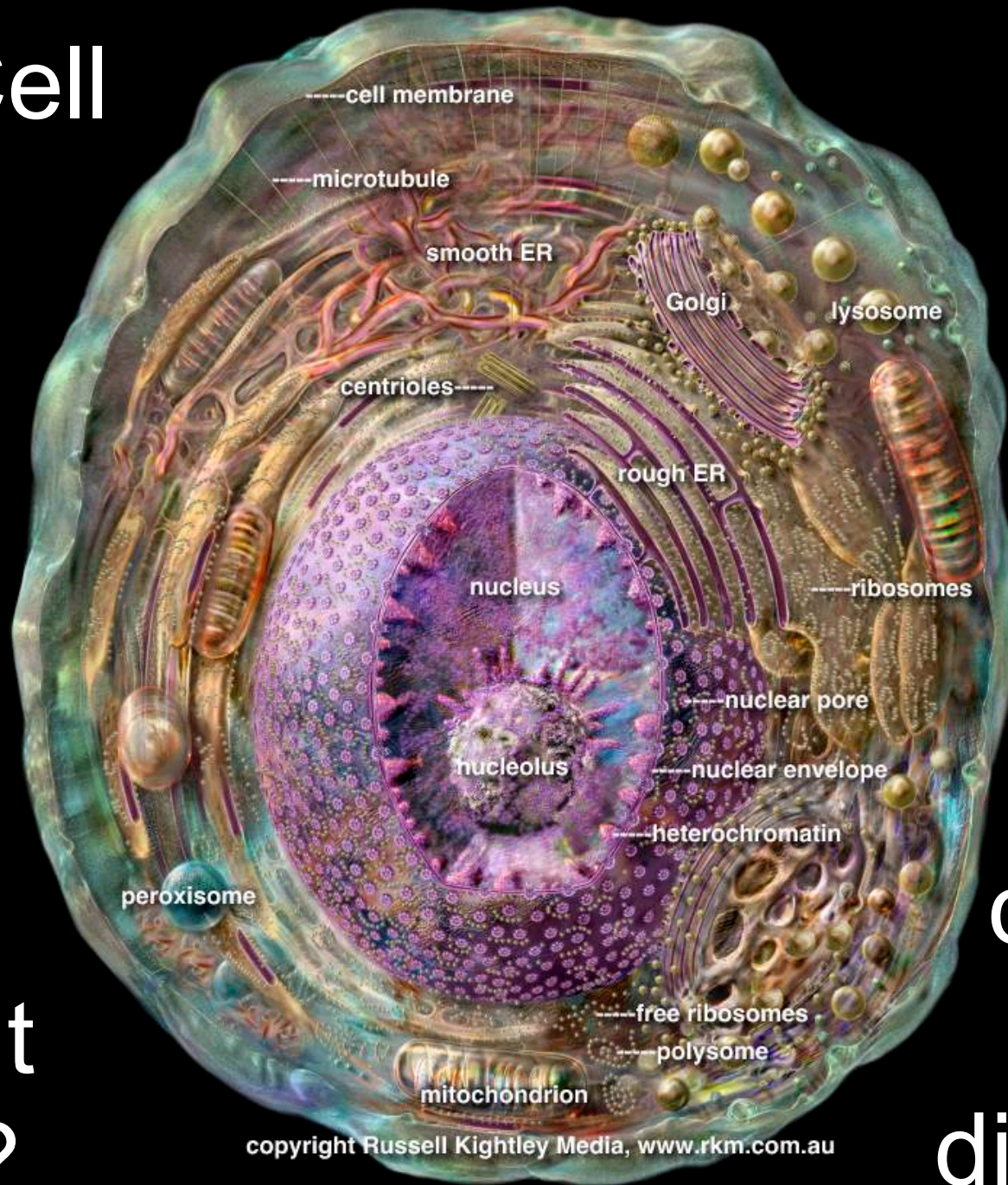
<http://baderlab.org>



Donnelly Center for Cellular and Biomolecular Research



The Cell



How
does it
work?

How
does it
fail in
disease?

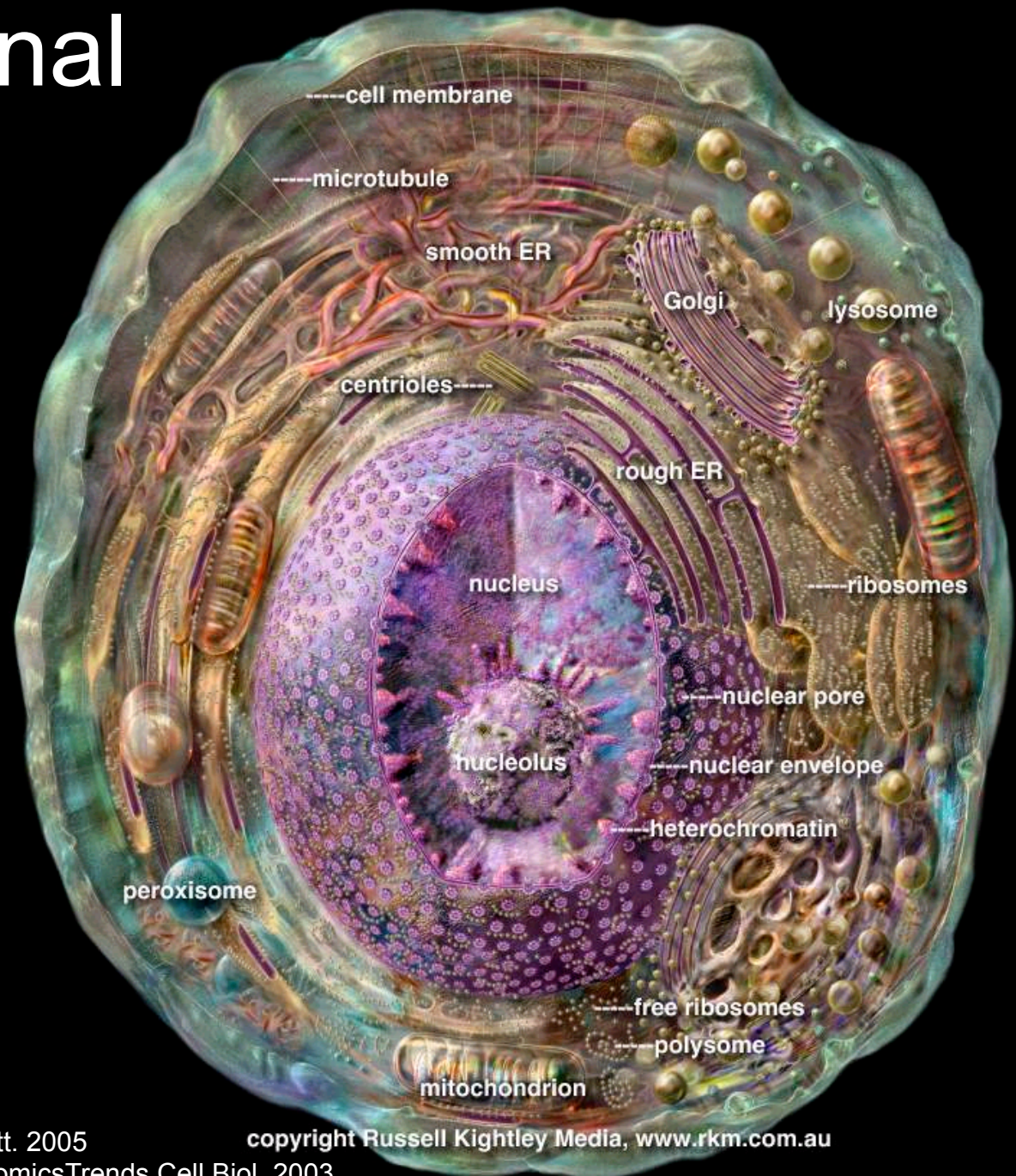
Computational Cell Map

Map the cell

- Predict map from genome
- Active cell map
- Multiple perturbation mapping
- Map visualization and analysis software

Read map to understand

- Cell processes
- Gene function
- Disease effects
- Map evolution



Cary MP et al. Pathway information... FEBS Lett. 2005

Bader GD et al. Functional genomics and proteomics Trends Cell Biol. 2003

copyright Russell Kightley Media, www.rkm.com.au

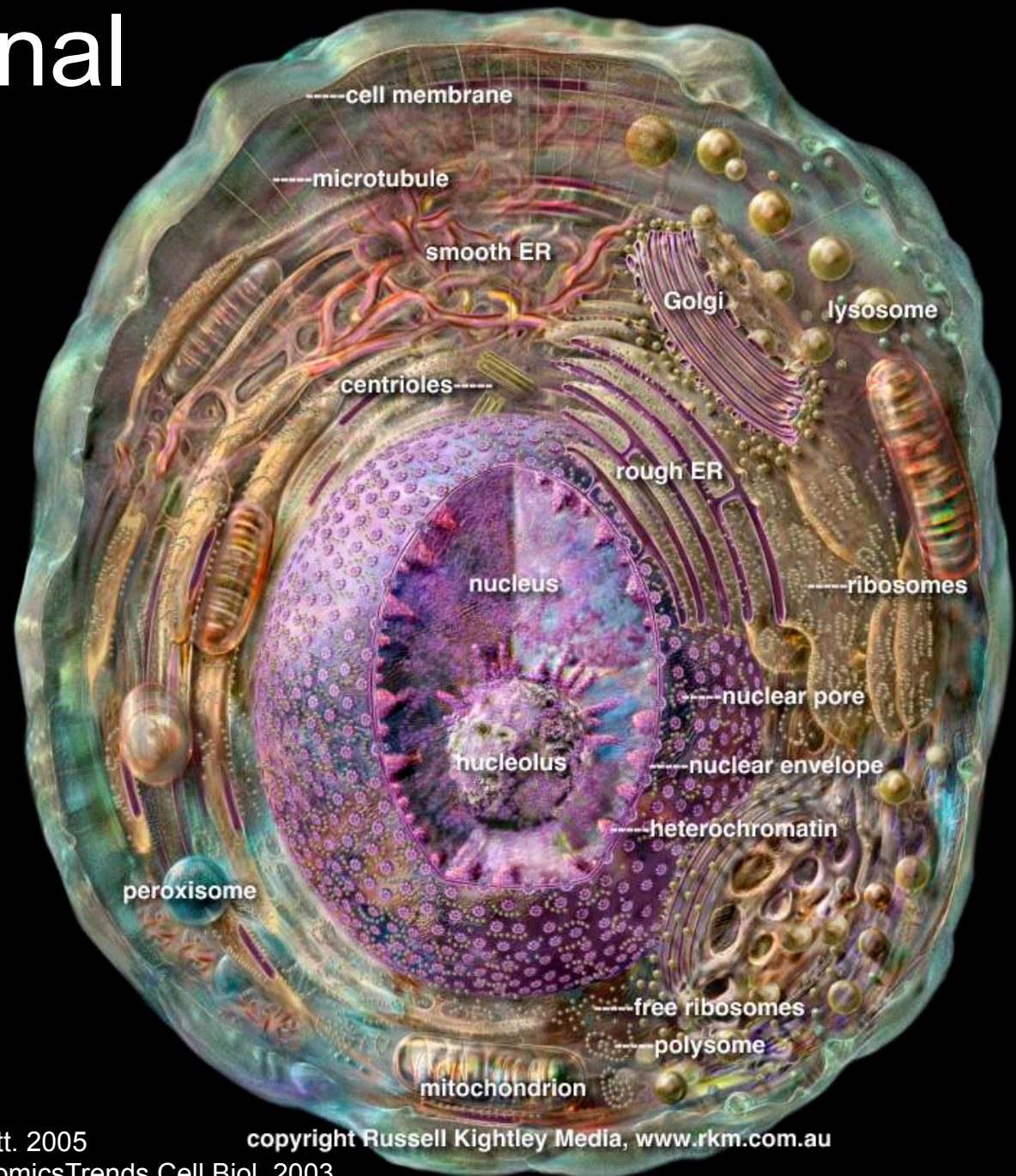
Computational Cell Map

Map the cell

- Predict map from genome
- Active cell map
- Multiple perturbation mapping
- Map visualization and analysis software

Read map to understand

- Cell processes
- Gene function
- Disease effects
- Map evolution

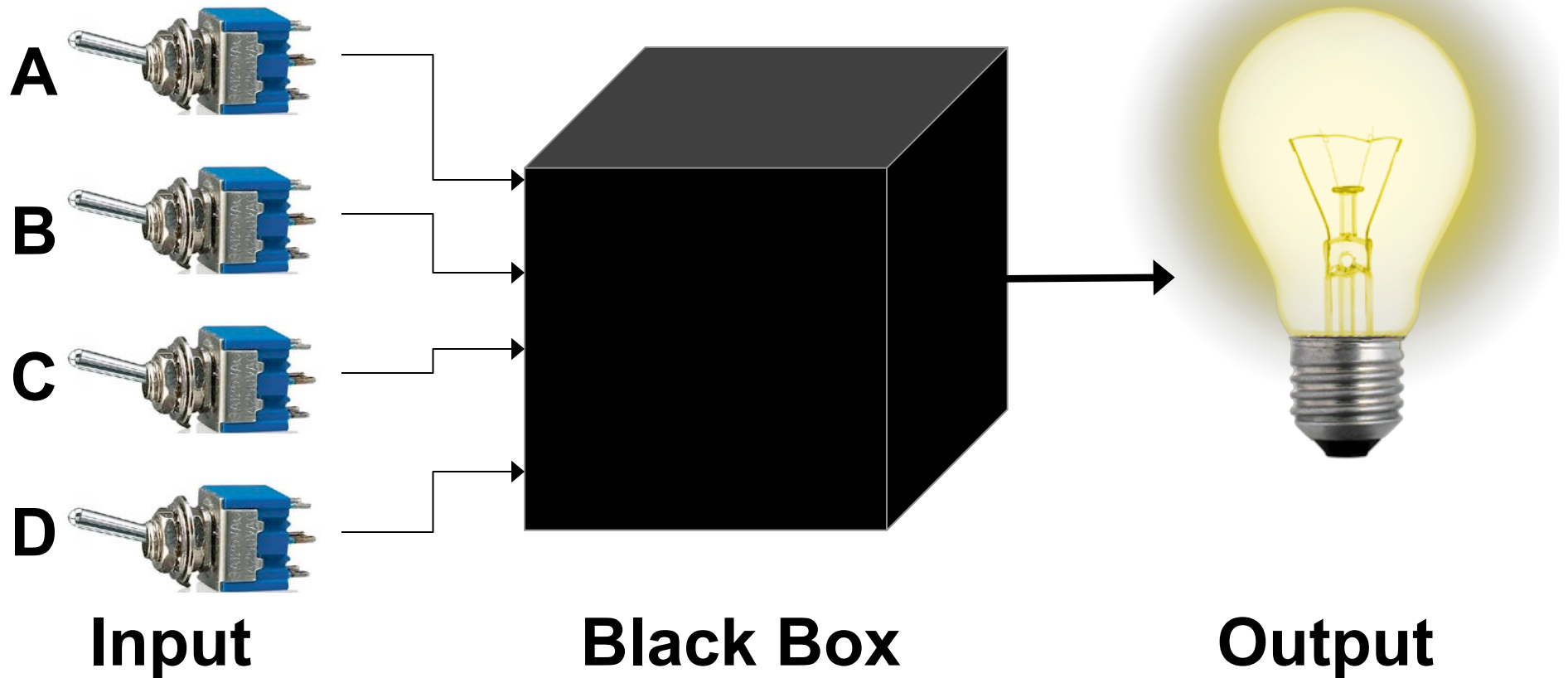


Cary MP et al. Pathway information... FEBS Lett. 2005

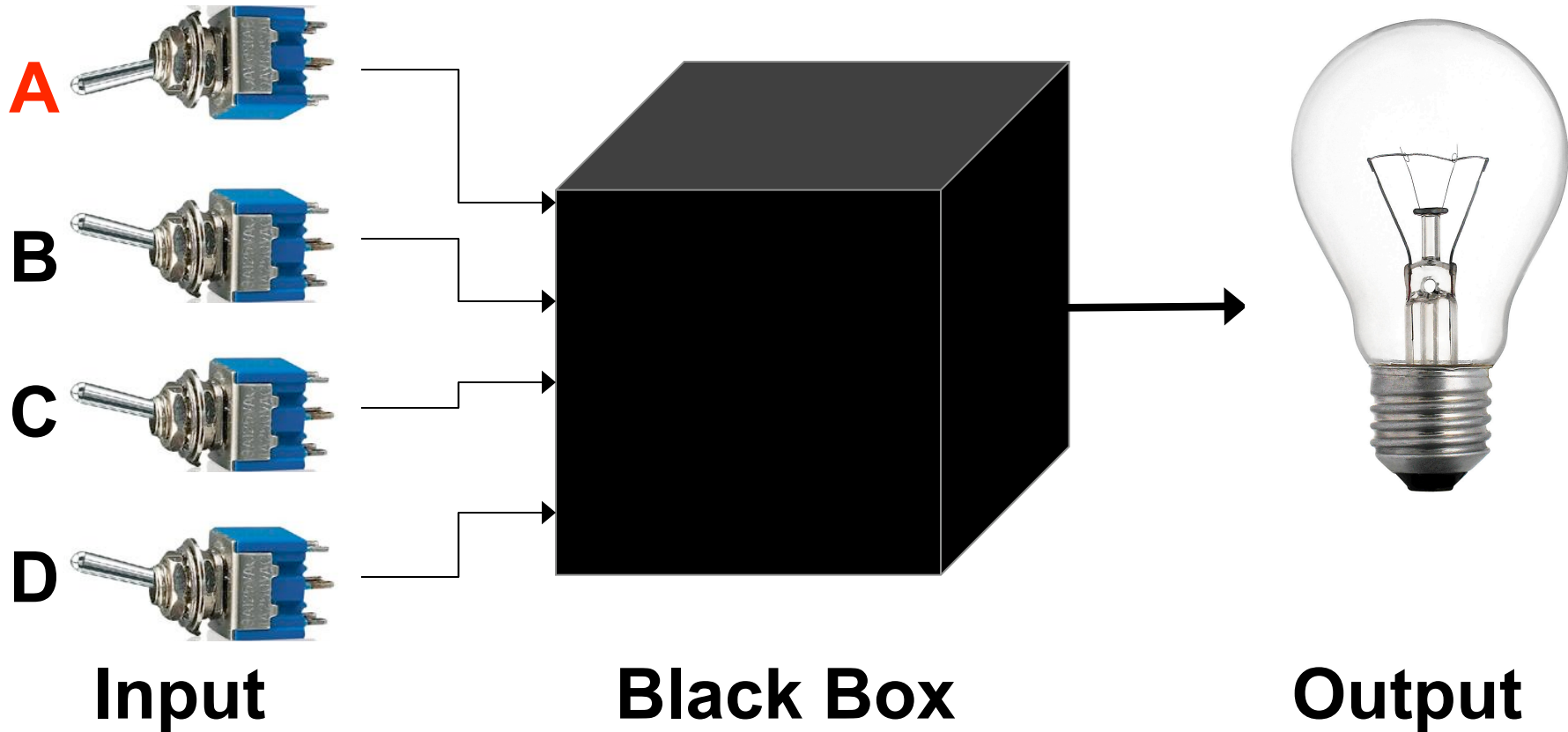
Bader GD et al. Functional genomics and proteomics Trends Cell Biol. 2003

copyright Russell Kightley Media, www.rkm.com.au

Multiple Perturbation Mapping

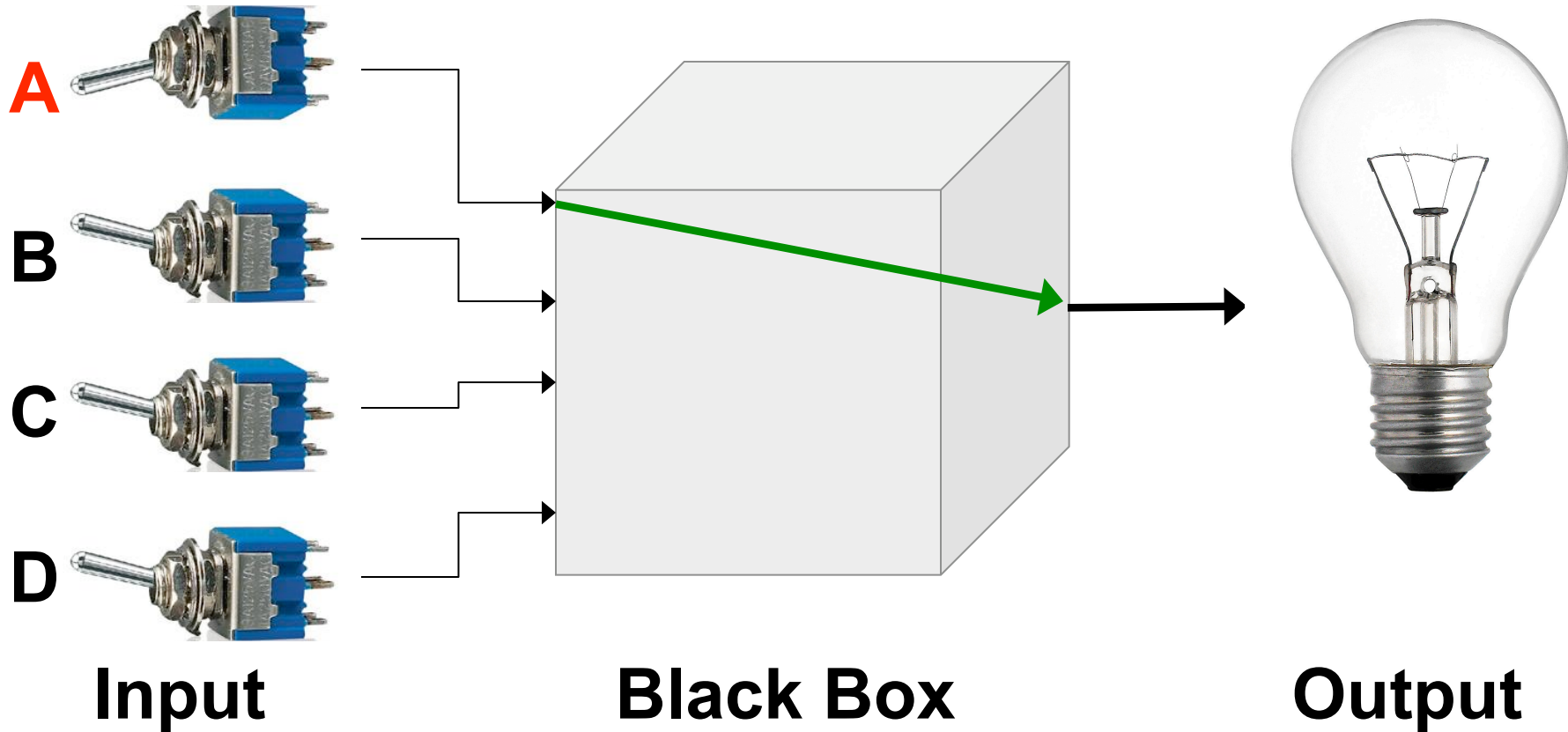


Direct Effect



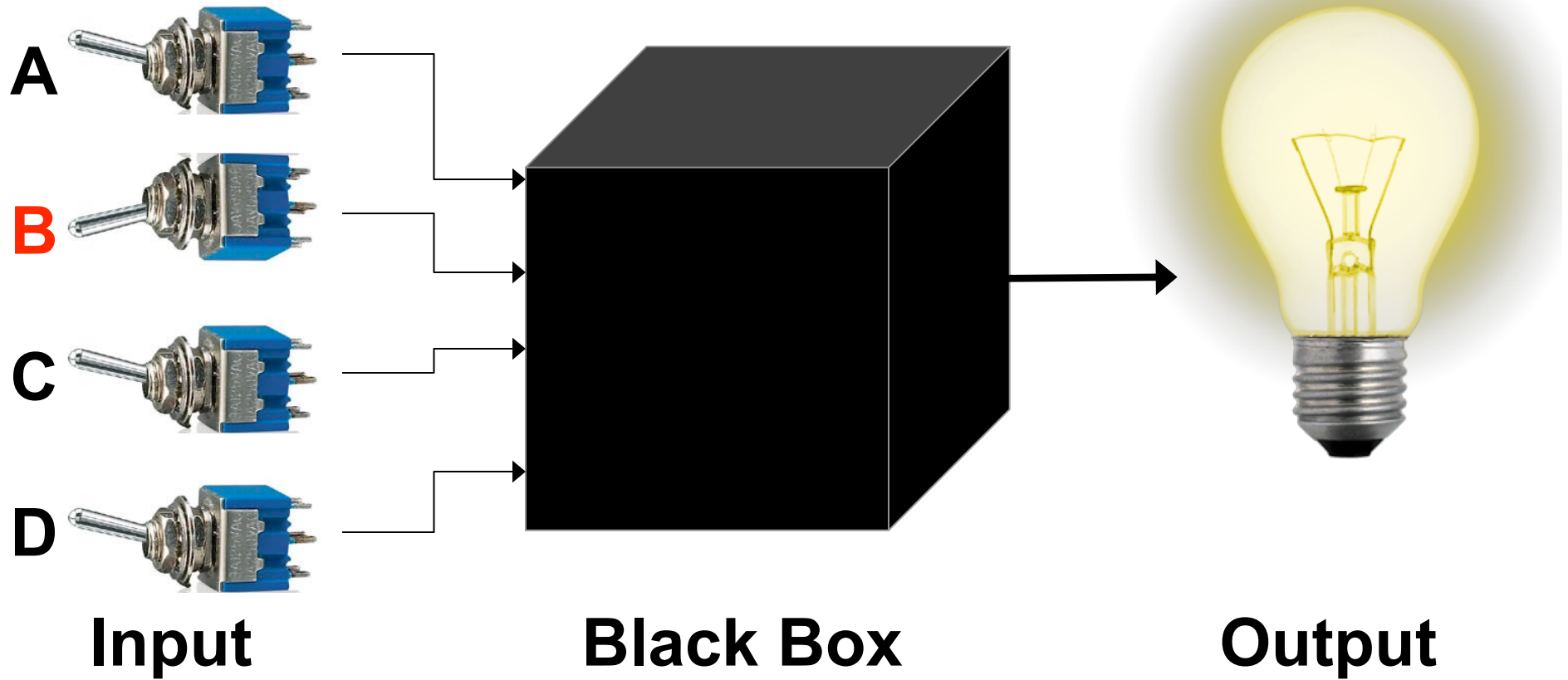
A Off → Light Off

Direct Effect

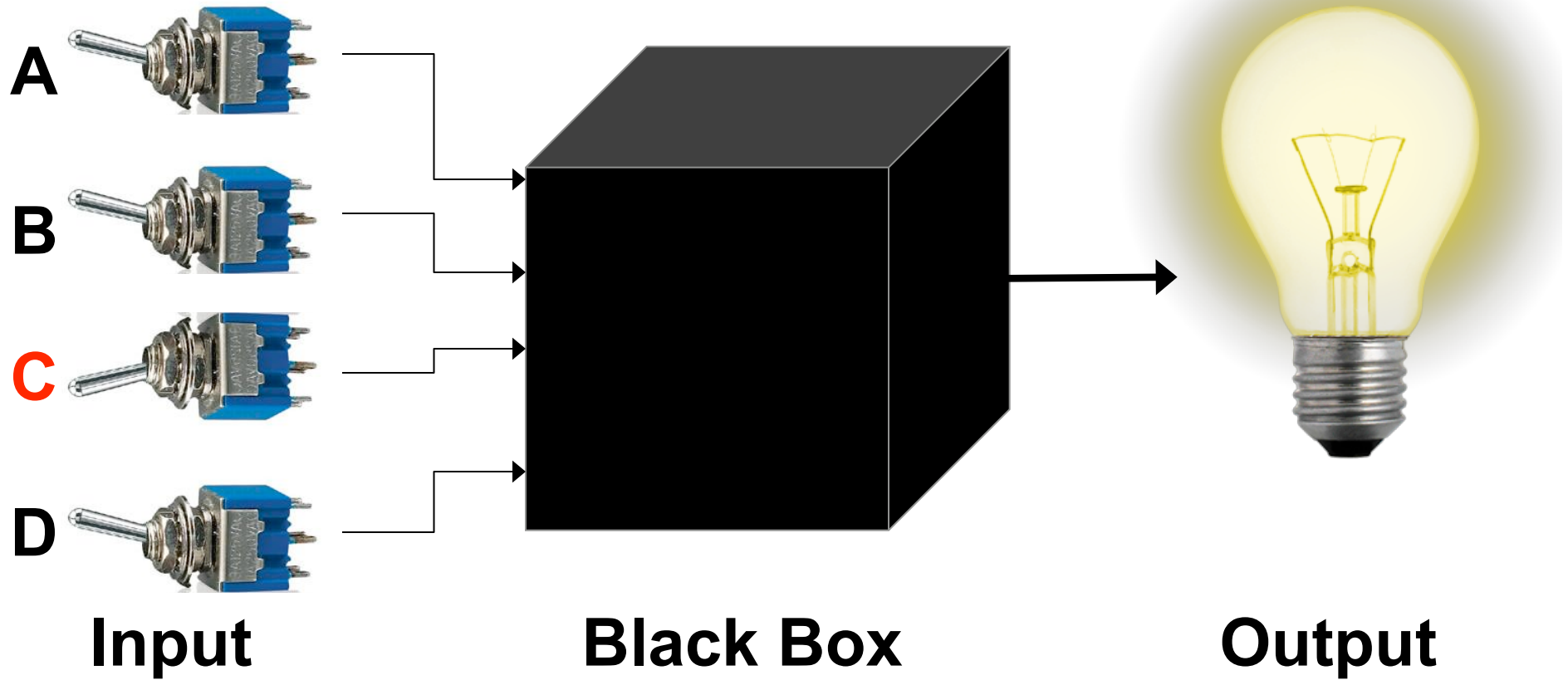


A Off → Light Off

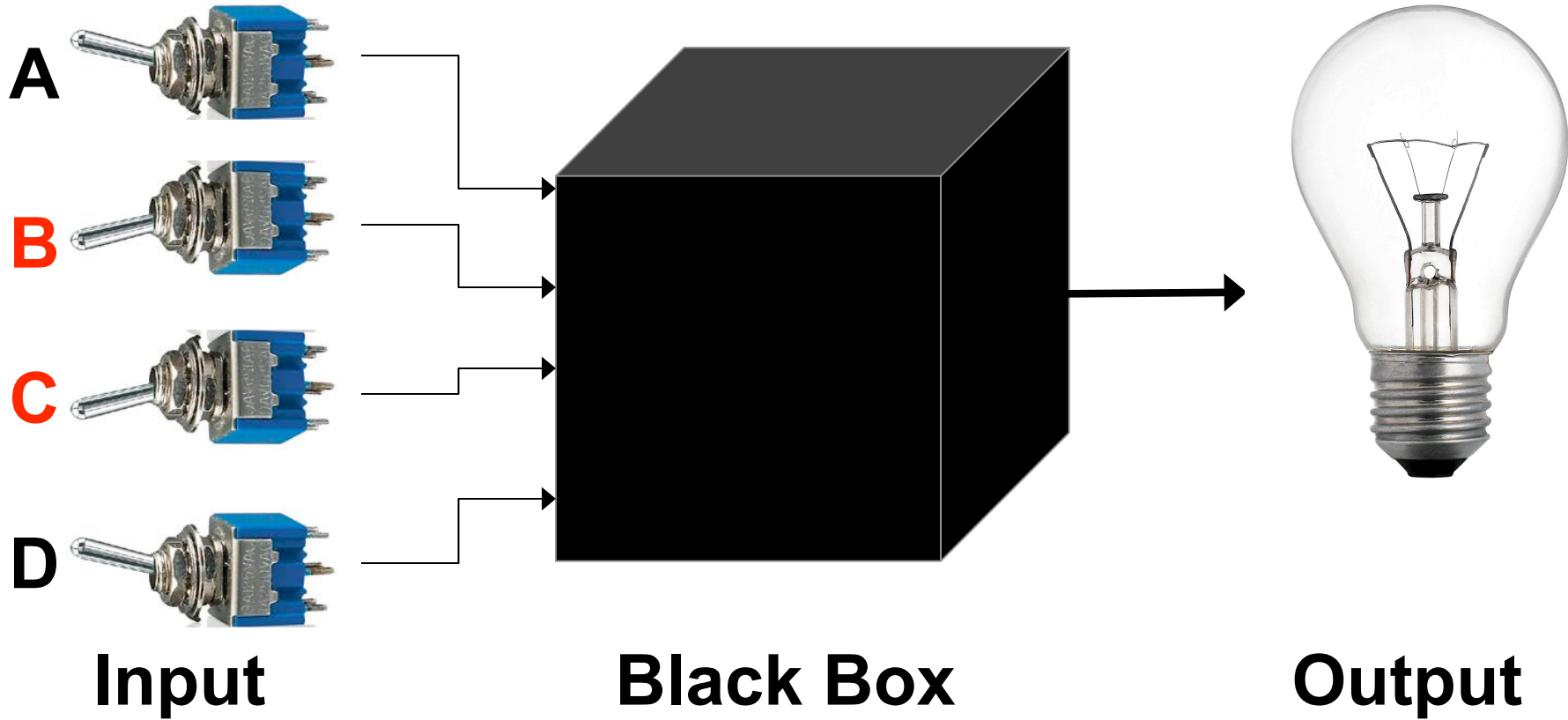
Parallel Paths



Parallel Paths

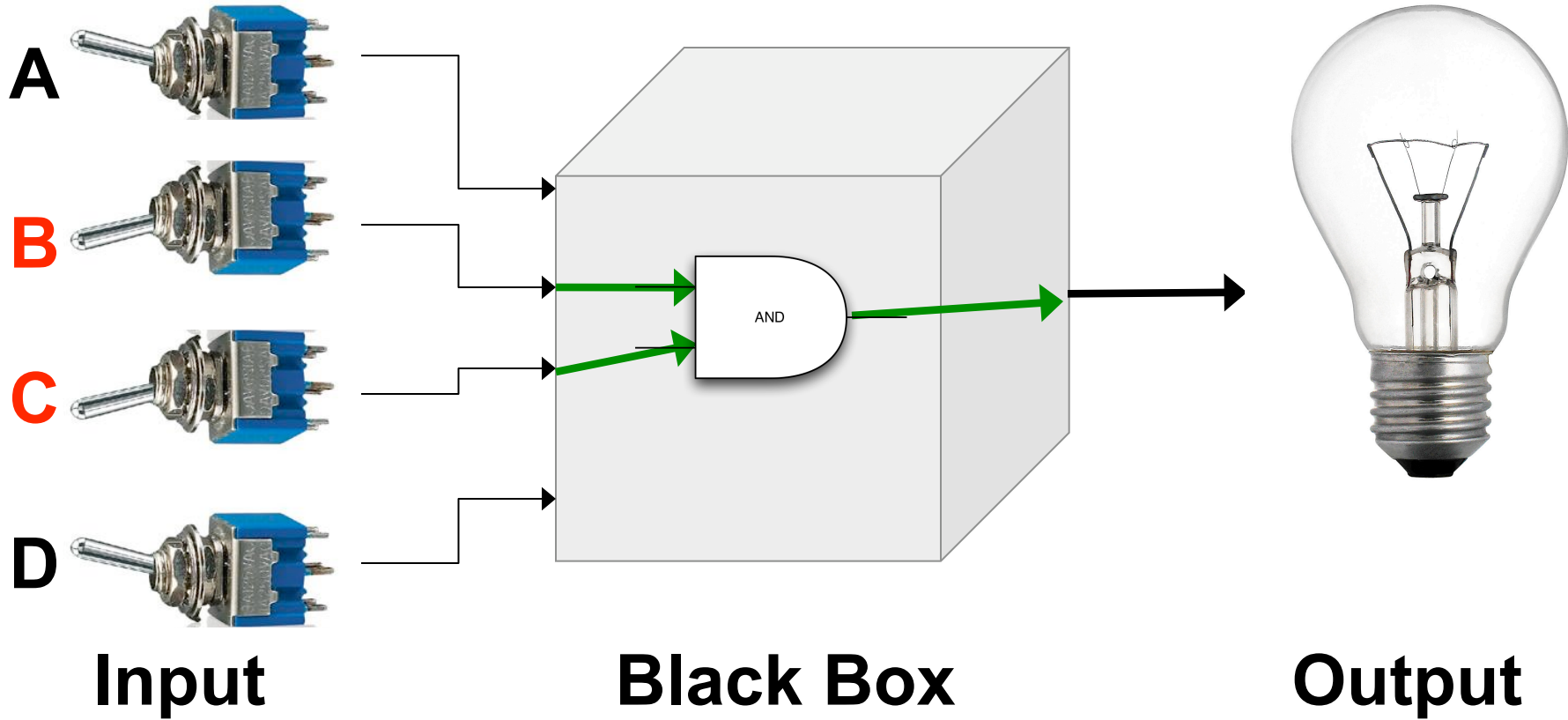


Parallel Paths



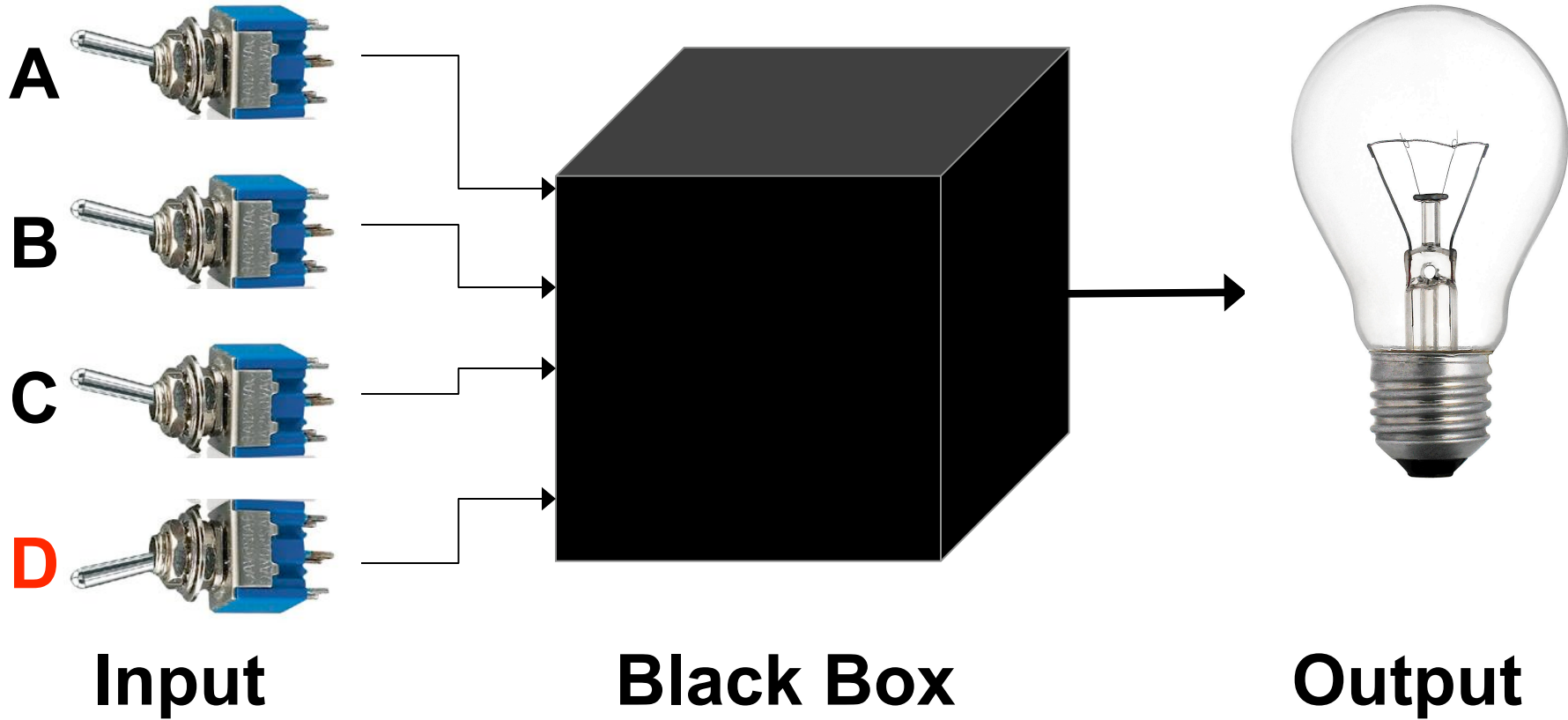
B&C Off → Light Off

Parallel Paths

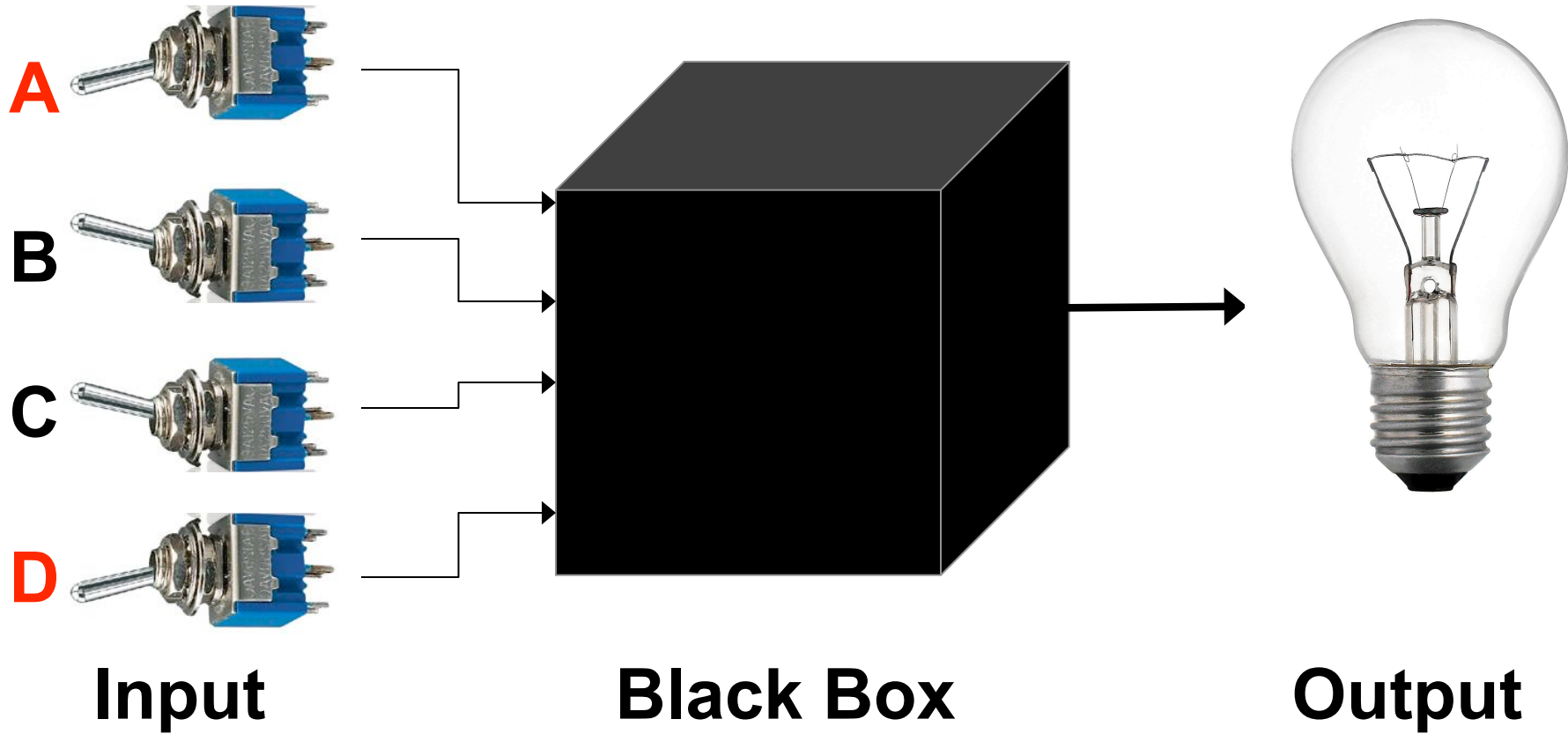


B&C Off → Light Off

Upstream Effect

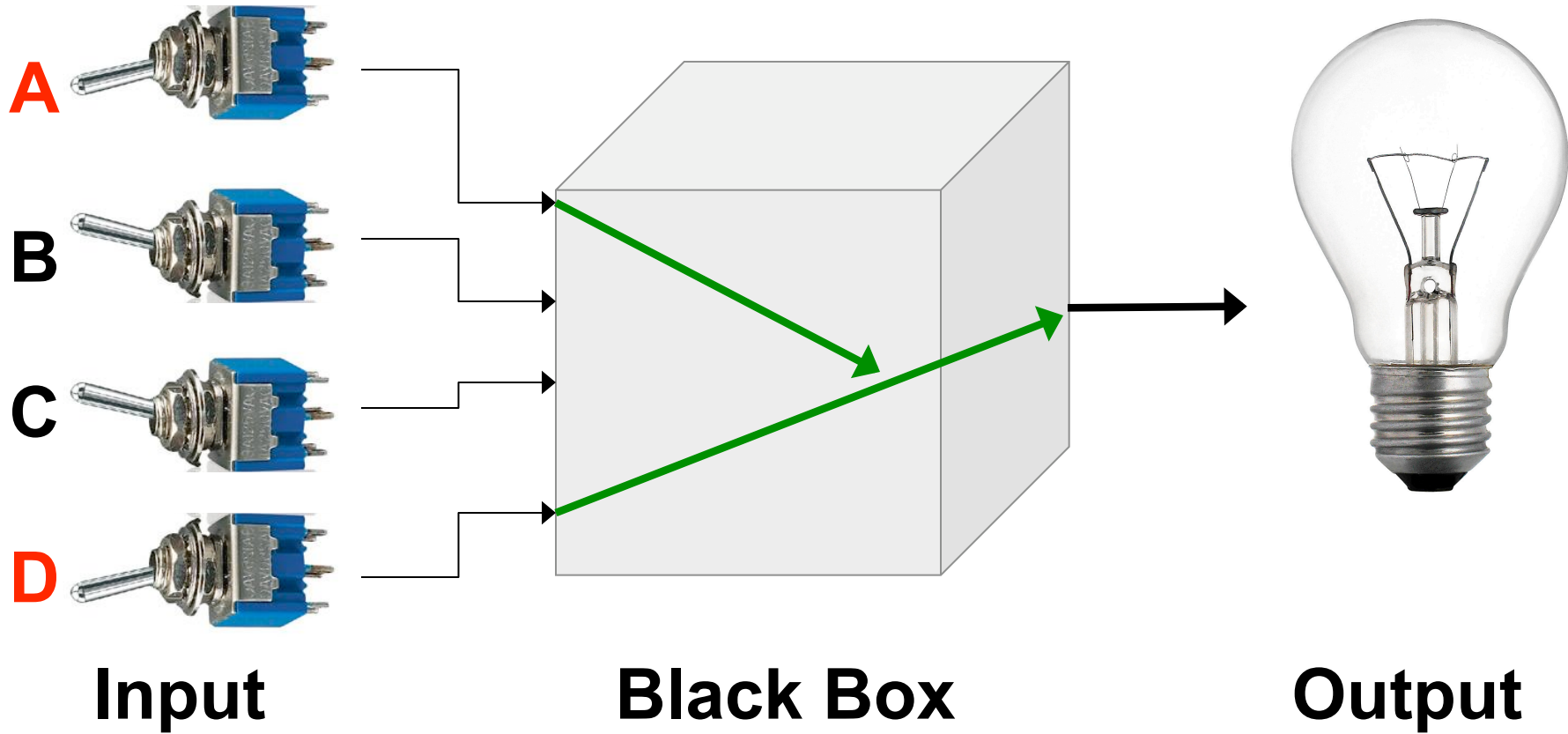


Upstream Effect



D masks A effect

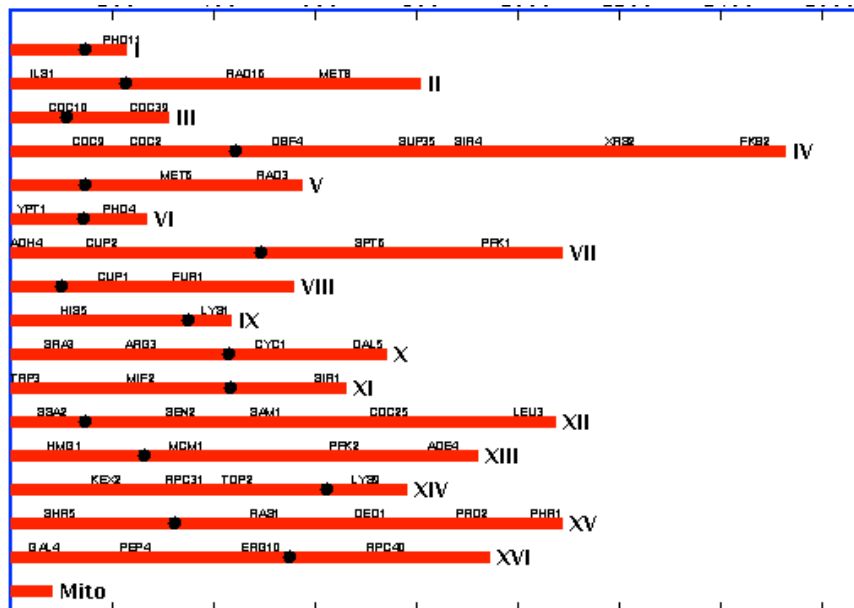
Upstream Effect



D masks A effect

Our Black Box: Budding Yeast

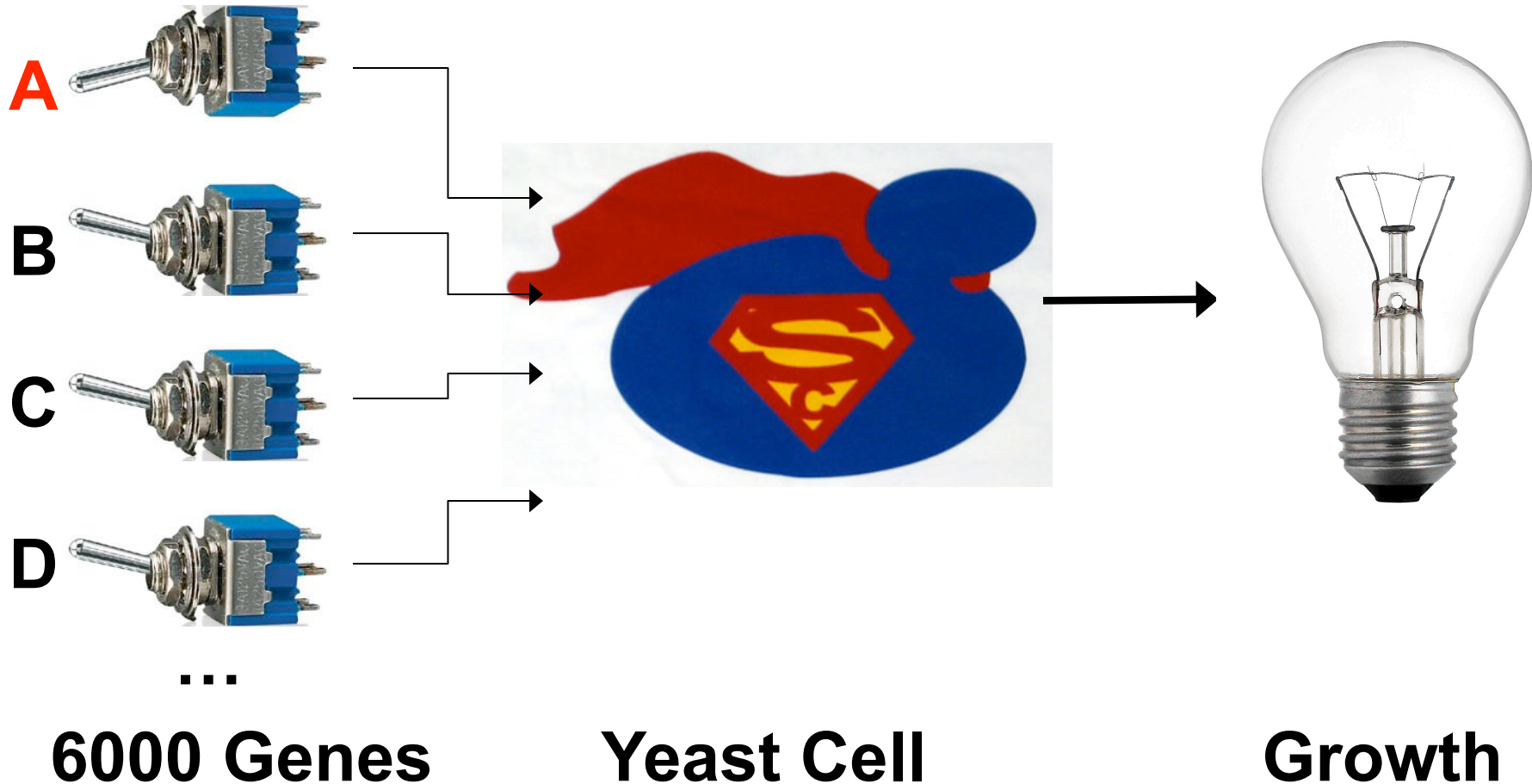
- a.k.a. Baker's Yeast, Brewer's Yeast
- *Saccharomyces cerevisiae* (Fungi)
 - Greek “sugar mold” + Latin “of beer”
- >6000 genes (parts)



Why Yeast?

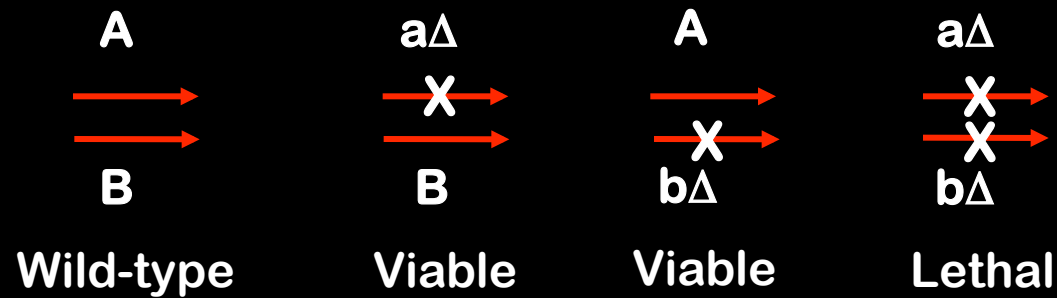
- Simple system (single cell)
- Easy to work with in a lab
- >20% yeast genes in human genome
 - Similar core processes
 - Nobel prize 2001: cell growth and cancer
- Pathogenic yeast, anti-fungal drugs
- We farm it: bread, beer, wine
 - Fermentation: sugar → ethanol, CO₂

Mapping Yeast

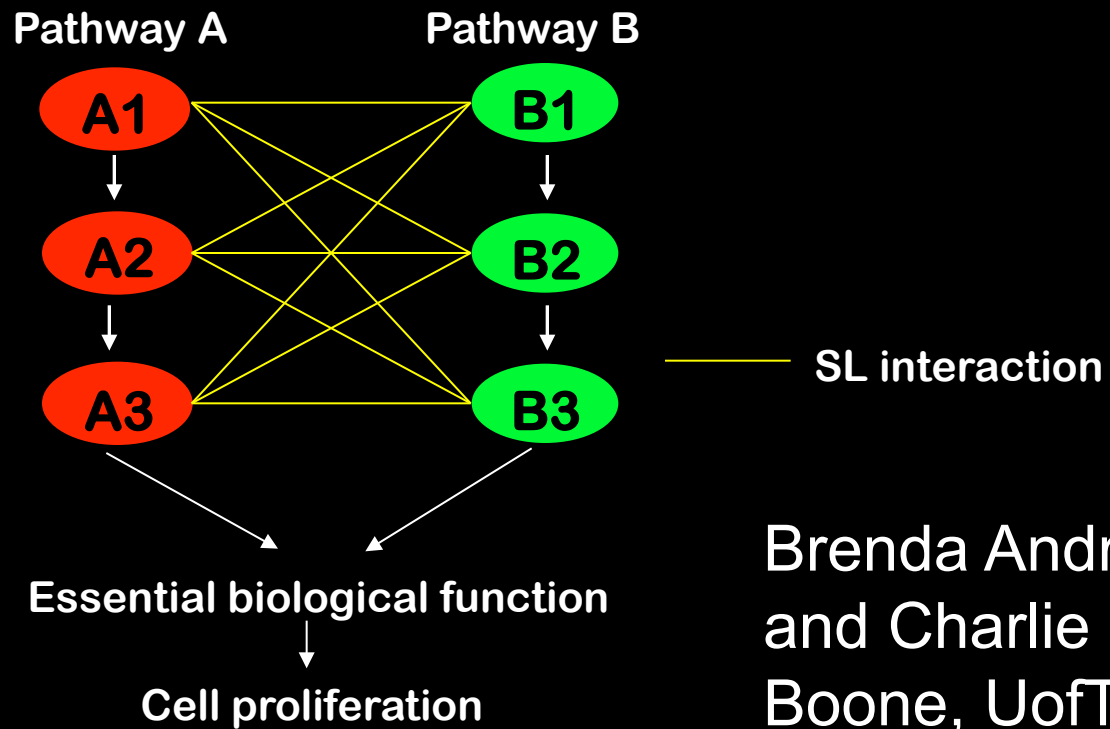
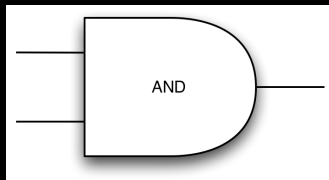


Single Gene Deletions: 1000 Genes Essential

Double Gene Deletions

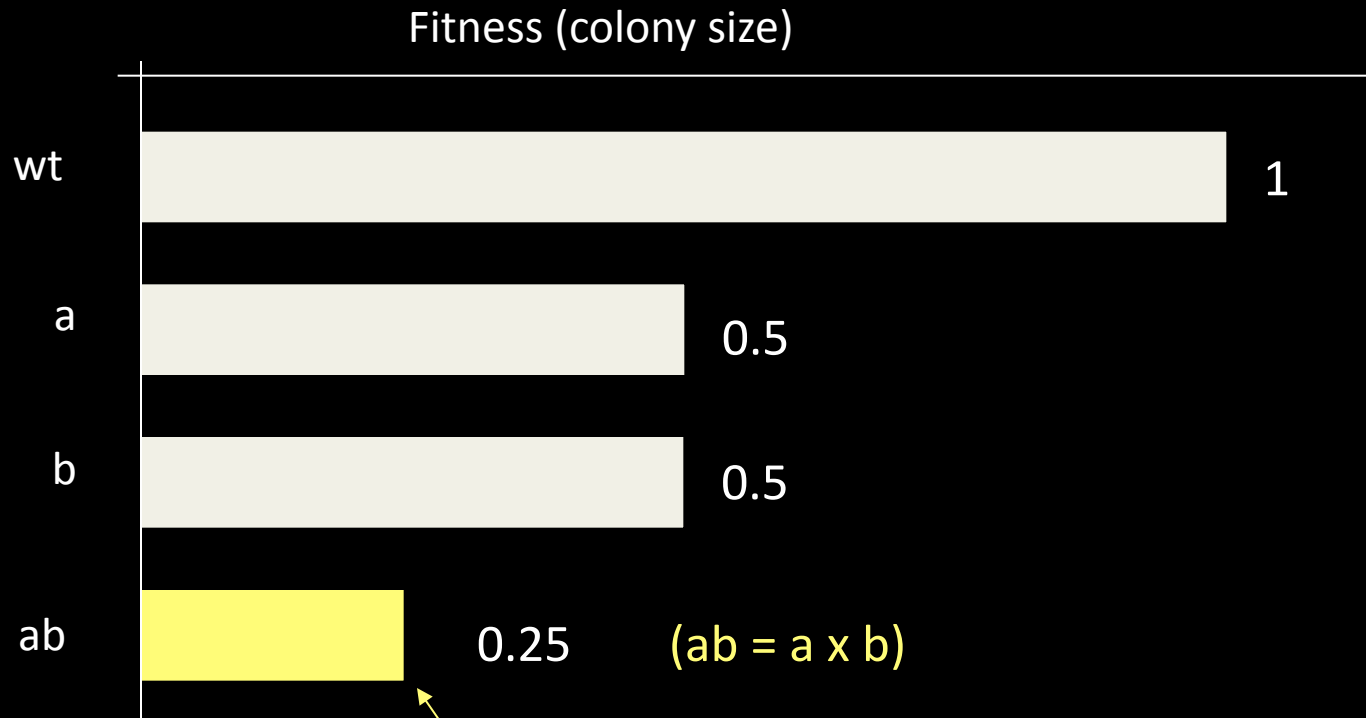


Genetic Interactions



Brenda Andrews
and Charlie
Boone, UofT

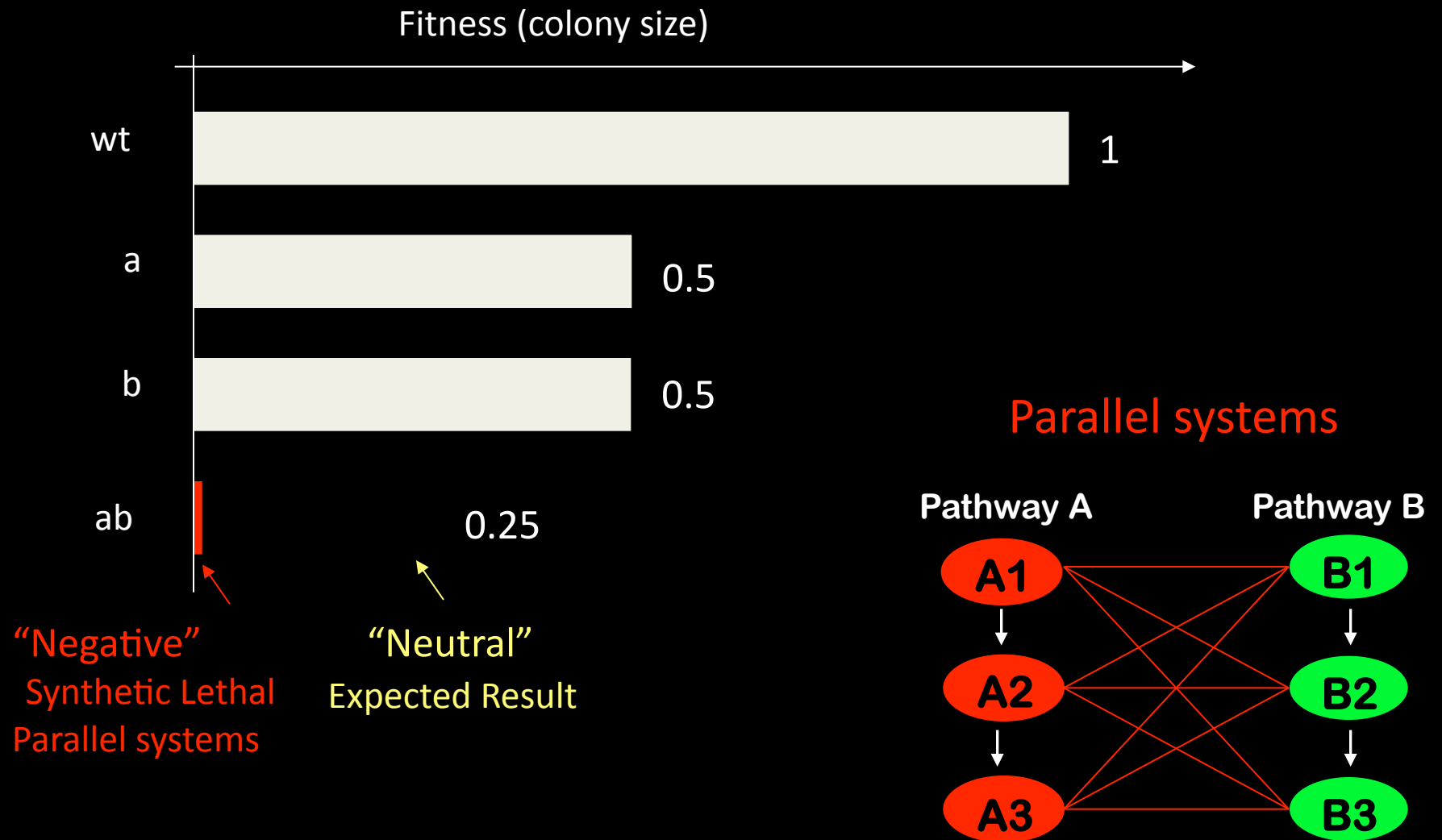
Three Basic Types of Genetic Interactions



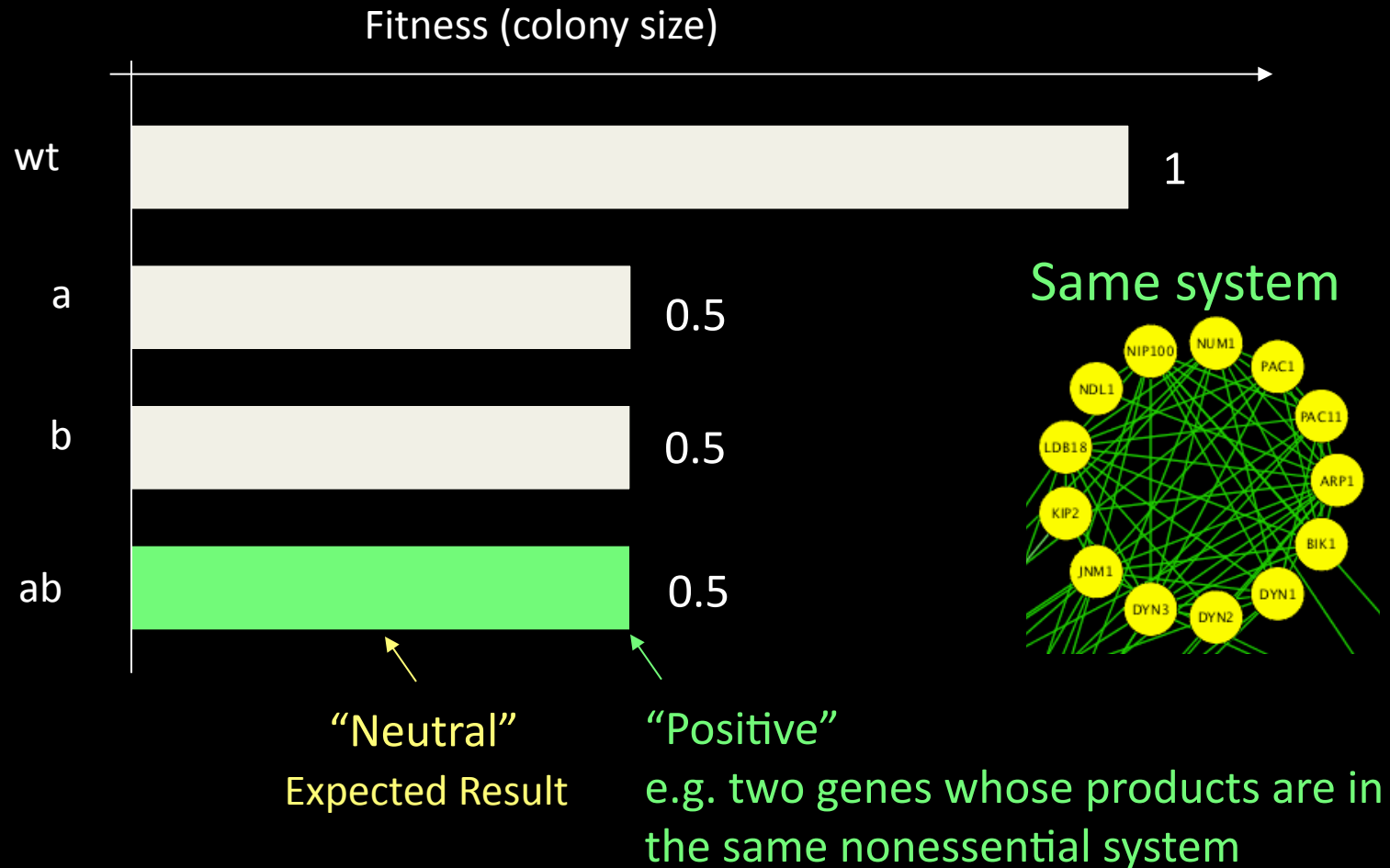
“Neutral”
Expected Result
Multiplicative Model



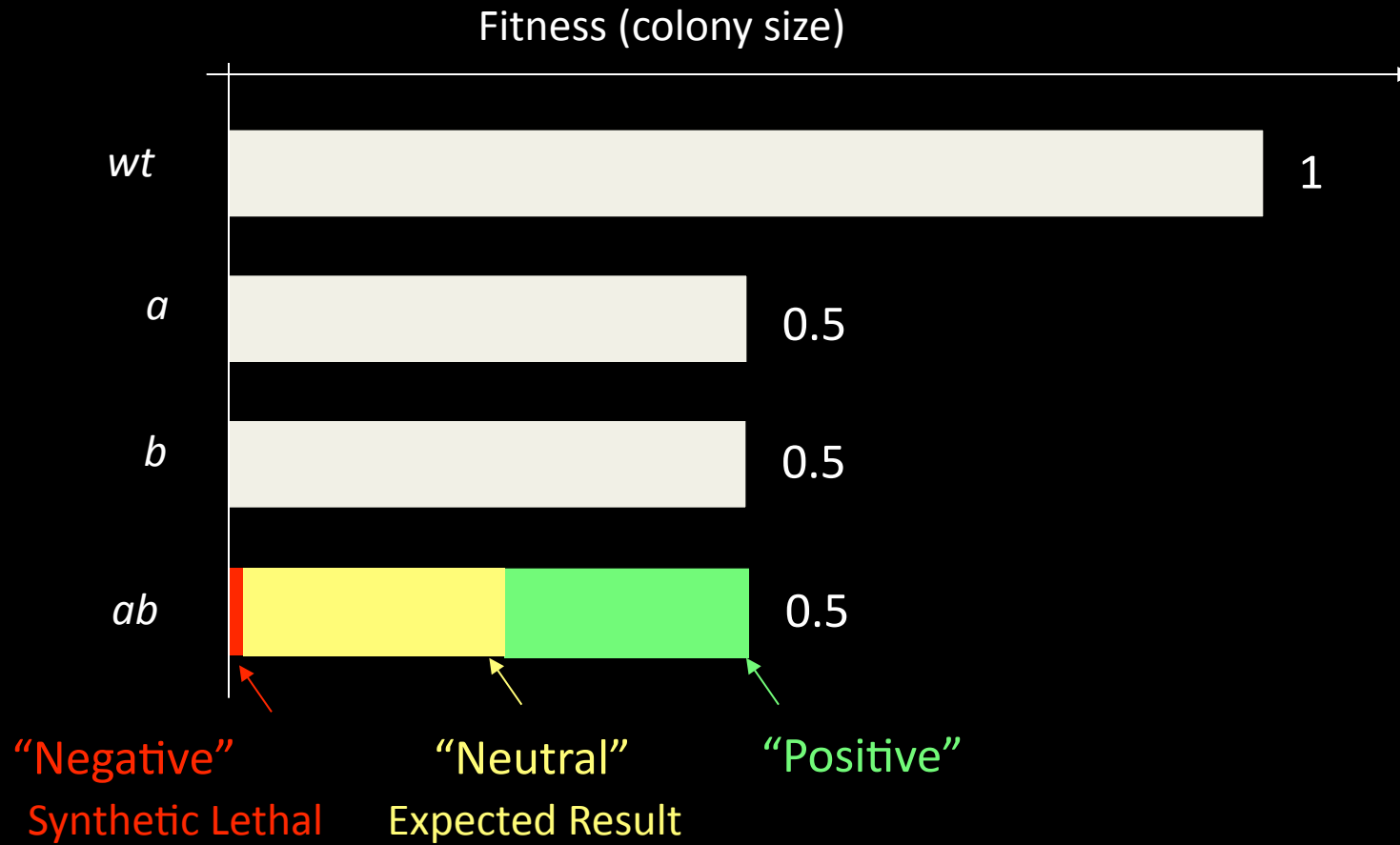
Three Basic Types of Genetic Interactions



Three Basic Types of Genetic Interactions



Three Basic Types of Genetic Interactions



Large-scale Mapping of Genetic Interactions in Yeast

6000 Yeast Genes:

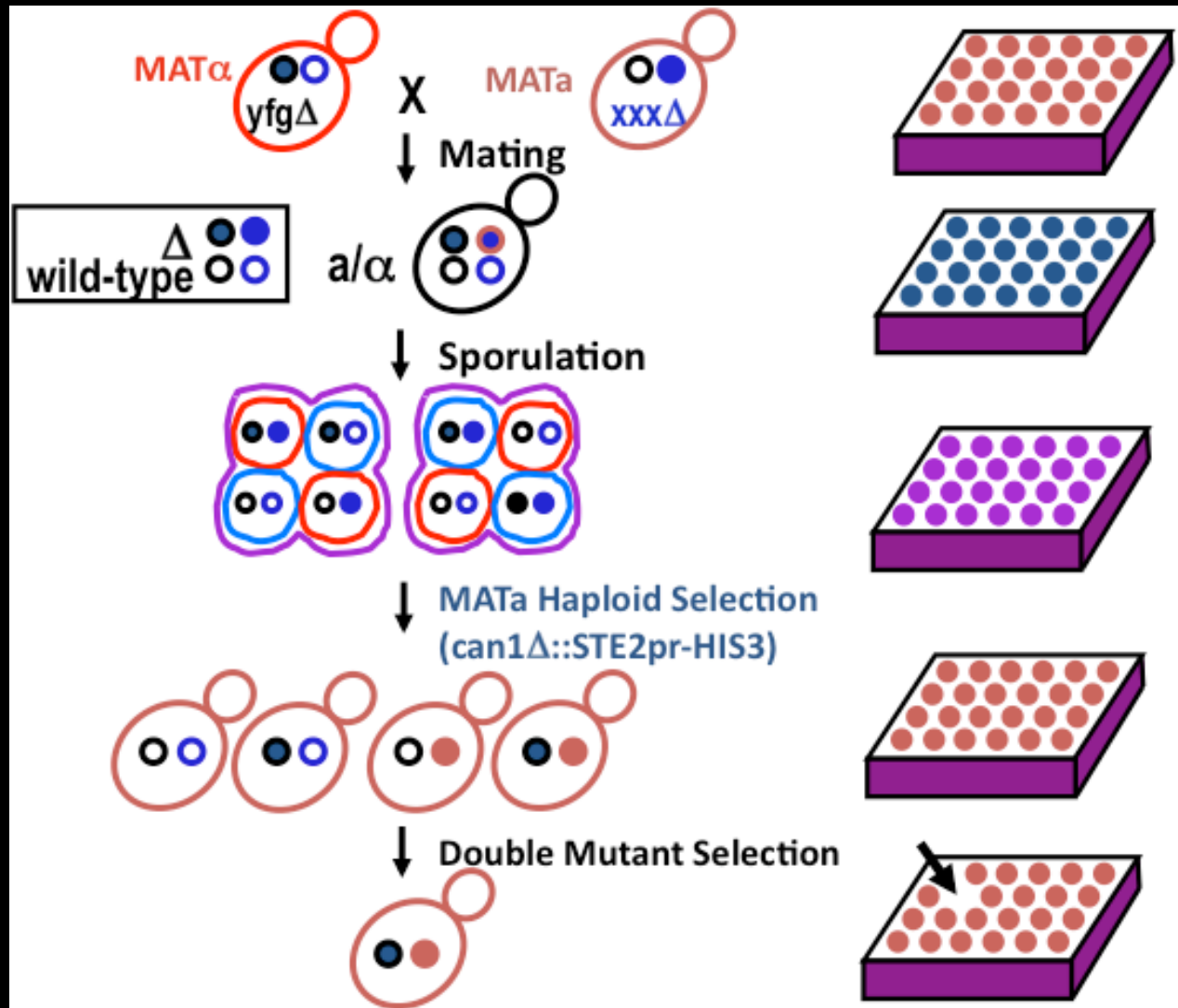
1000 Essential Genes

5000 Nonessential Genes

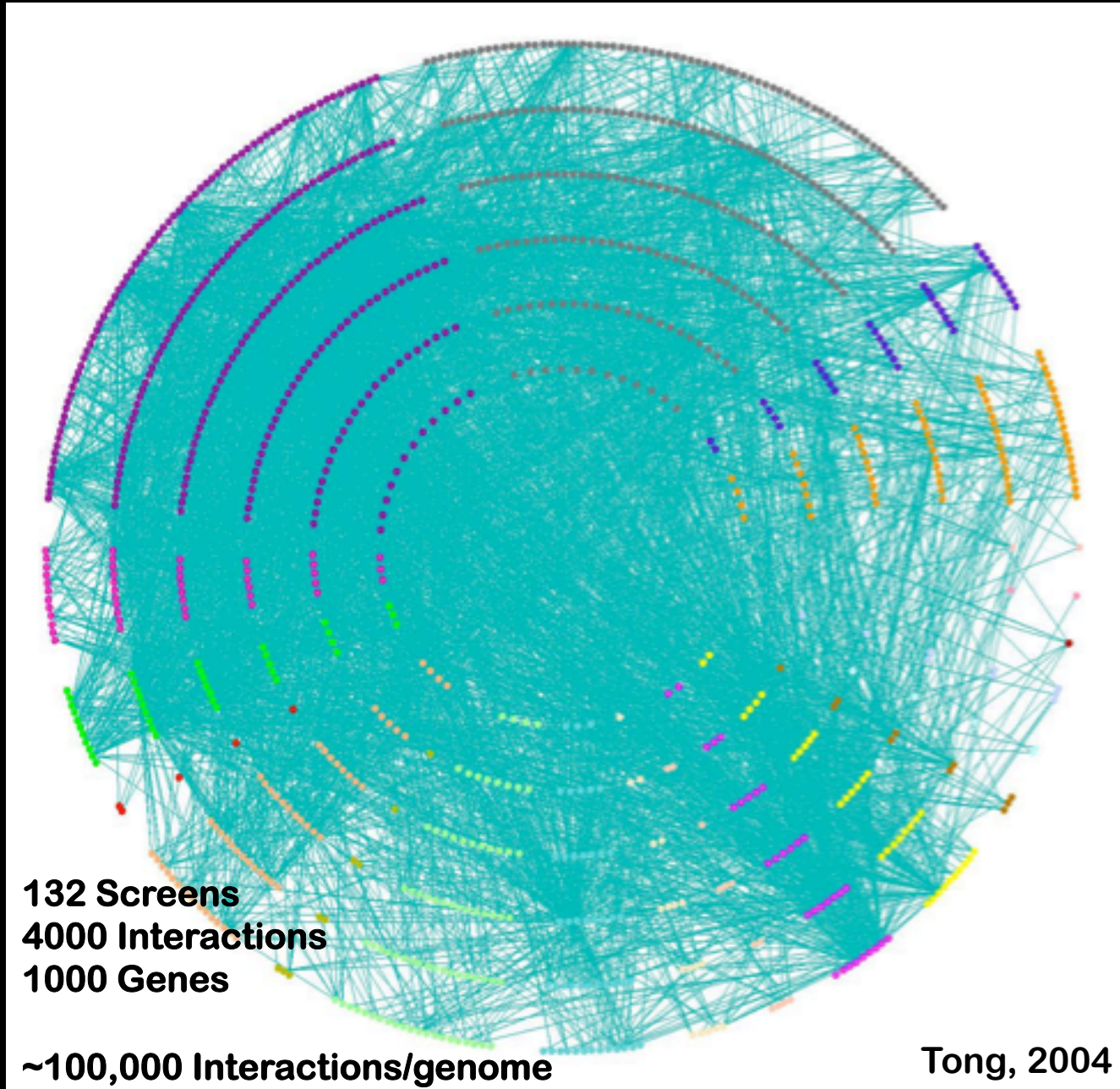
- Genetic Array ~ 5000 Viable Yeast Deletion Mutants
~ 1000 conditional alleles of essential genes
- Automated Genetics
- Examine 36 Million Double Mutants & Map GIs

Brenda Andrews
and Charlie
Boone, UofT

Synthetic Genetic Array (SGA)

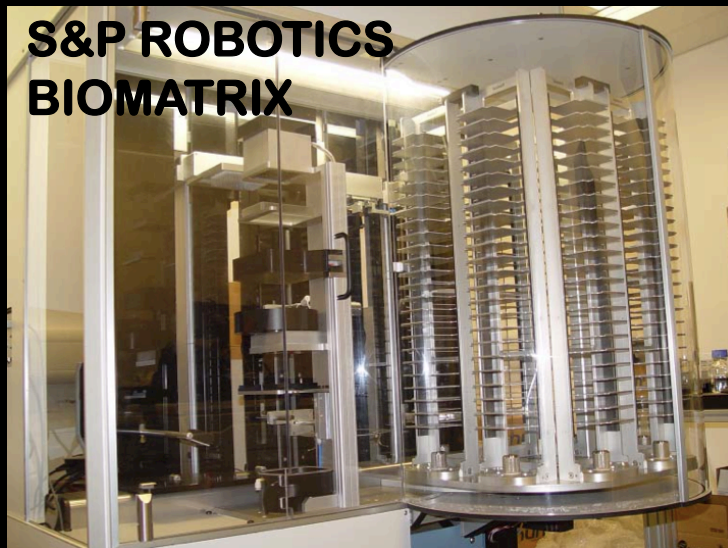
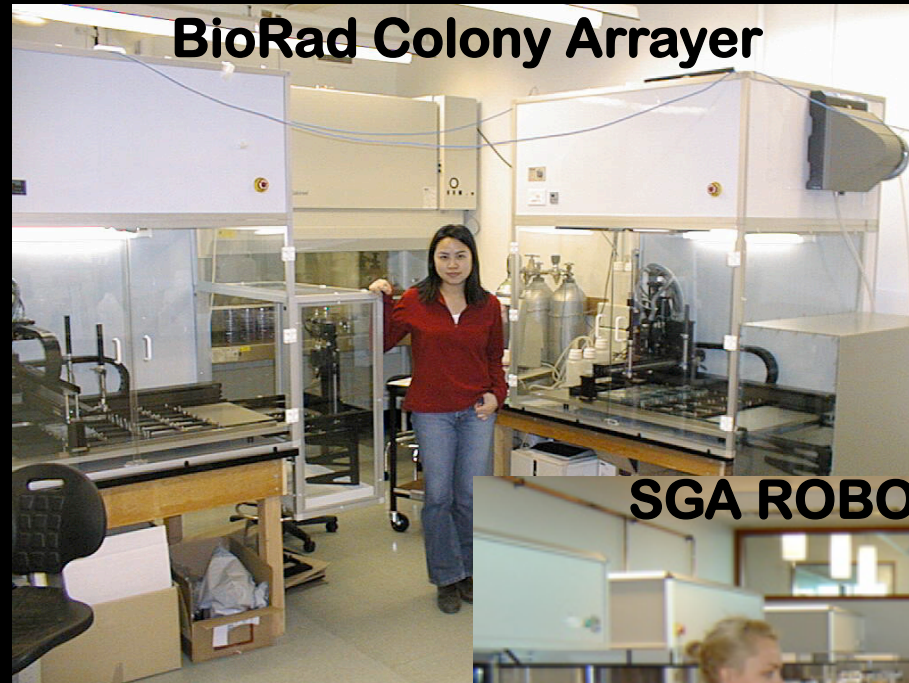


Large-Scale Genetic Interaction Network





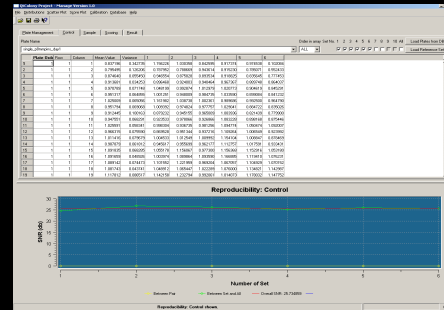
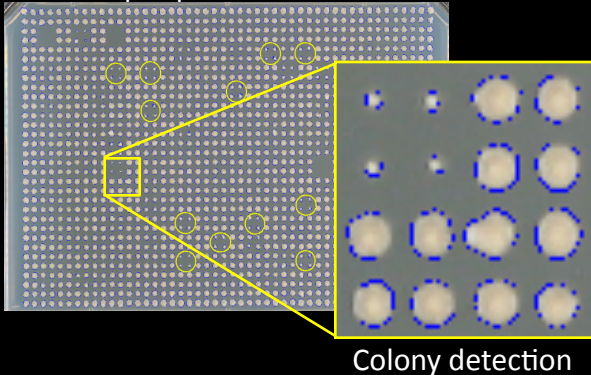
Automated Yeast Genetics



200 genome-wide screens/month

Quantitative Genetic Interactions from Double Mutant Colony Growth Modeling

SGA output plate



Model: double mutant colony size is multiplicative in *biological* and *experimental* factors

$$C_{ij} = \alpha f_i f_j \varepsilon_{ij} t \text{ pos}(C_{ij}) \text{ comp}(C_{ij}) \text{ batch}(C_{ij}) \dots e$$

Double Mutant Colony Size

Single Mutant Fitness

Genetic Interaction

Time

Experimental Factors

E.g. Plate Position
Nutrient Competition
Screen Batch

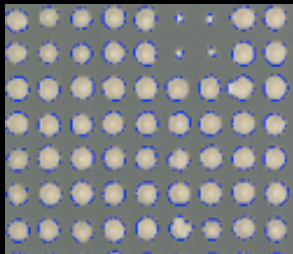
Log-normal error

Challenge: Most of the variation in colony sizes is due to systematic experimental factors, which we can estimate and normalize out, not biological effects

Fitting the Model: Deriving Precise Measures of Double and Single Mutant Fitness

$$C_{ij} = \alpha f_i f_j \varepsilon_{ij} t \text{ pos}(C_{ij}) \text{ comp}(C_{ij}) \text{ batch}(C_{ij}) \dots e$$

Double Mutant Colony Image



Experimental factor normalization

Camera adjustment

Plate normalization

Competition normalization

Spatial gradient normalization

Plate-specific row/col. normalization

Batch normalization

Cell # adjustment

Single mutant fitness estimation

f_i
 f_j

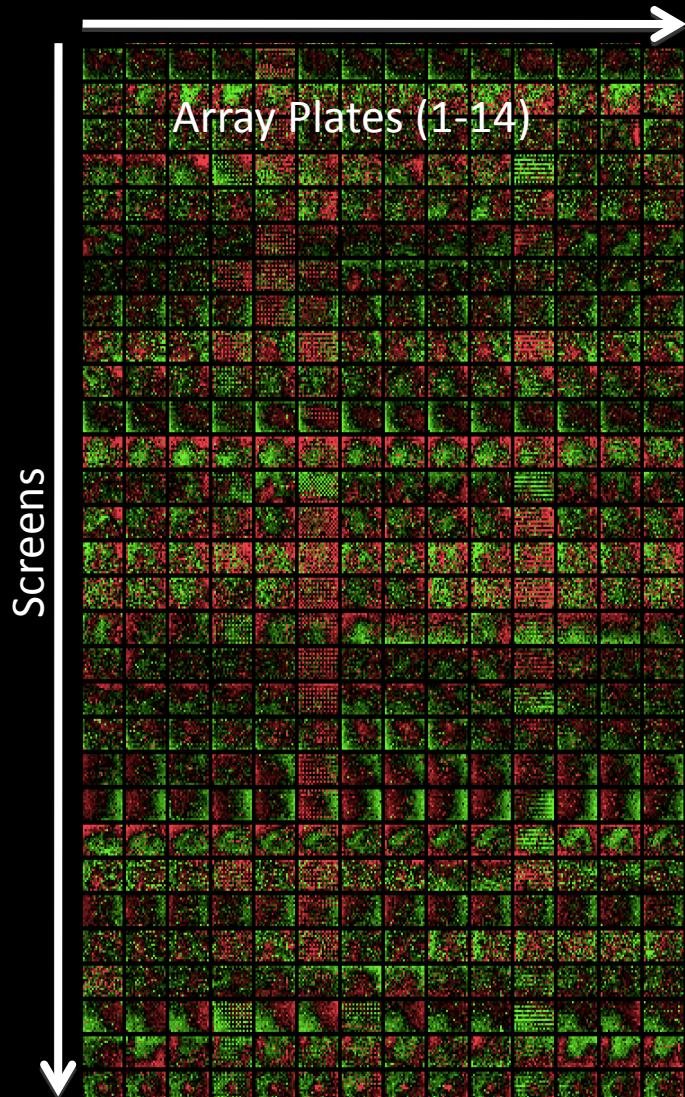
(with error estimates)

Genetic interaction/double mutant fitness estimation

f_{ij}, ε_{ij}

Normalizing Spatial and Batch Effects

Unnormalized interactions

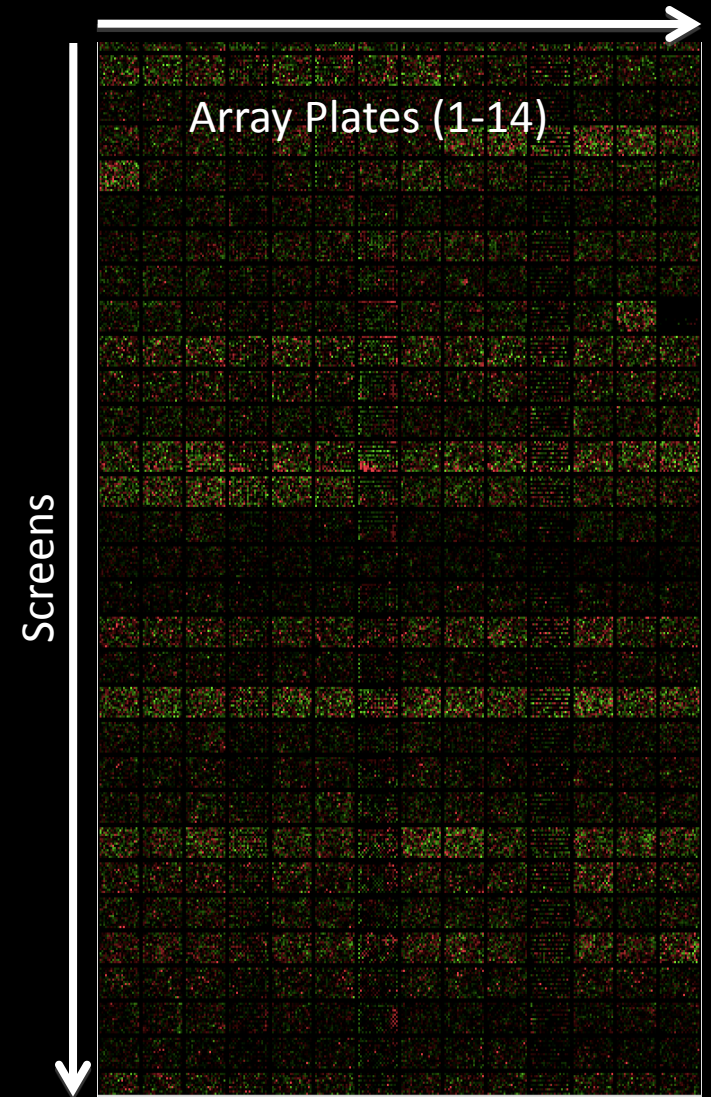


- Red square: Negative interaction (SL/SS)
- Green square: Positive interaction

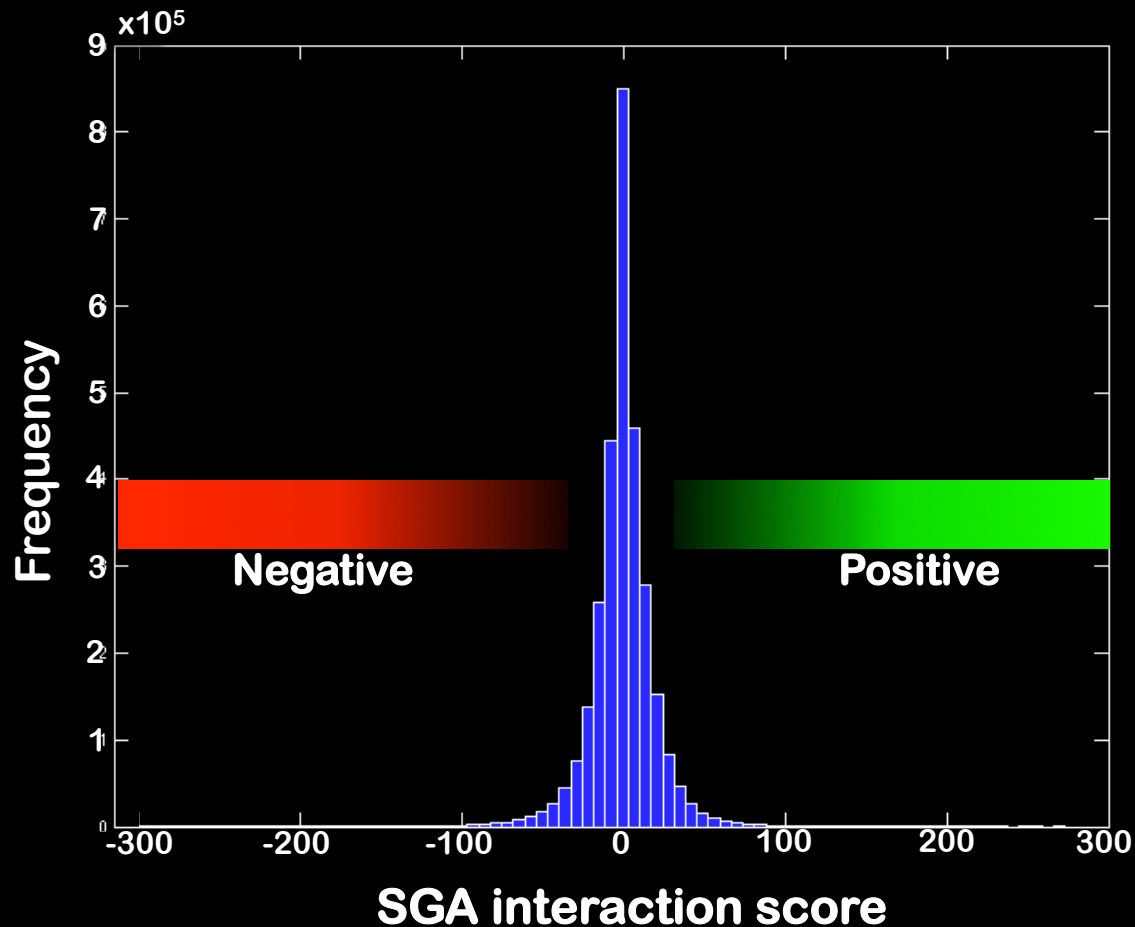
Correction



Normalized interactions



SGA Interaction Score Distribution



$\sim 3 \times 10^6$ gene pairs tested

At 95% confidence:

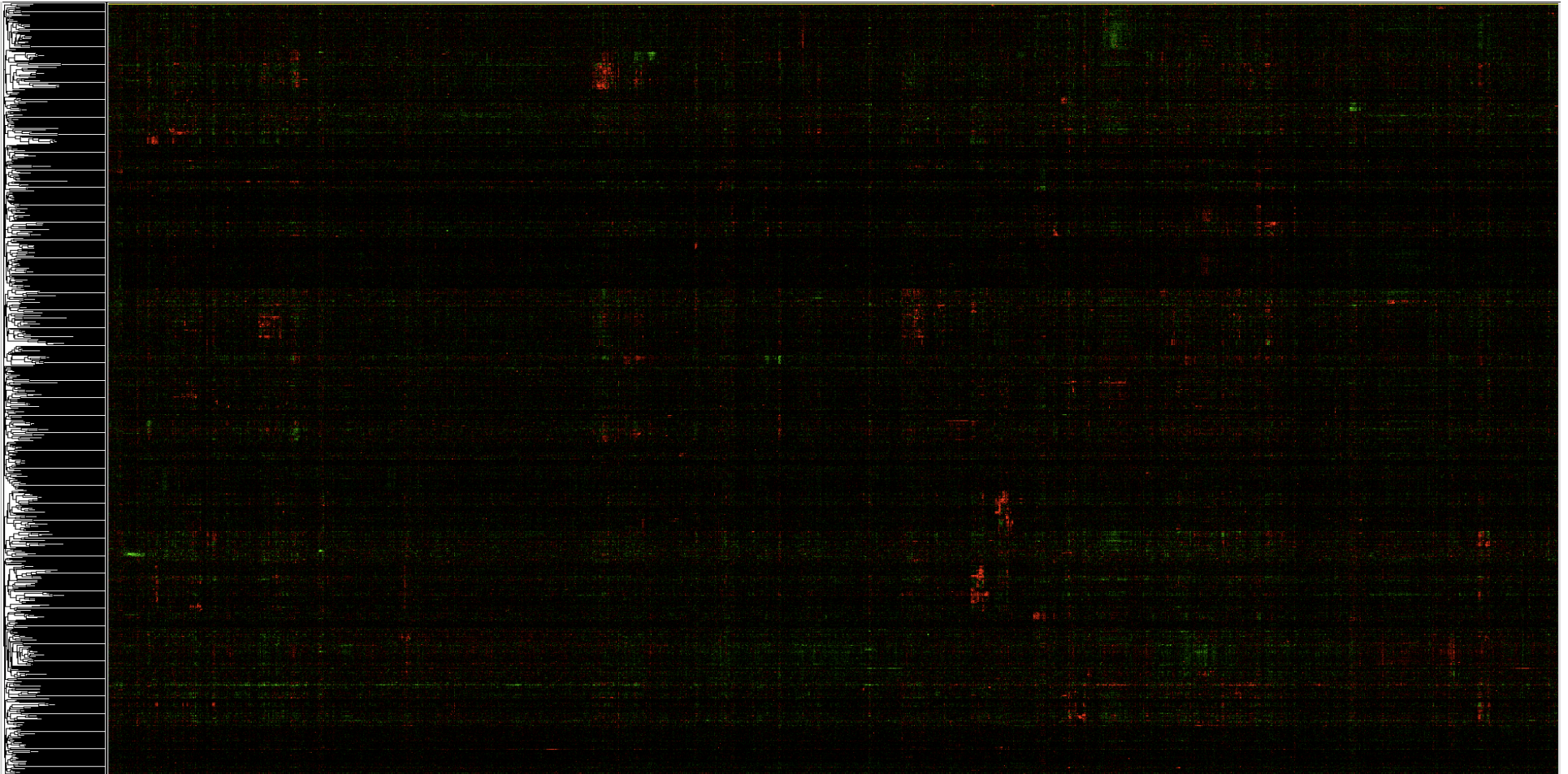
$\sim 45,000$ negative interactions

$\sim 20,000$ positive interactions

More negatives
than positives

Full genome SGA matrix

994 x ~5000



Quantitative Genetic Interaction Data Display Wall, Princeton

Chad Myers

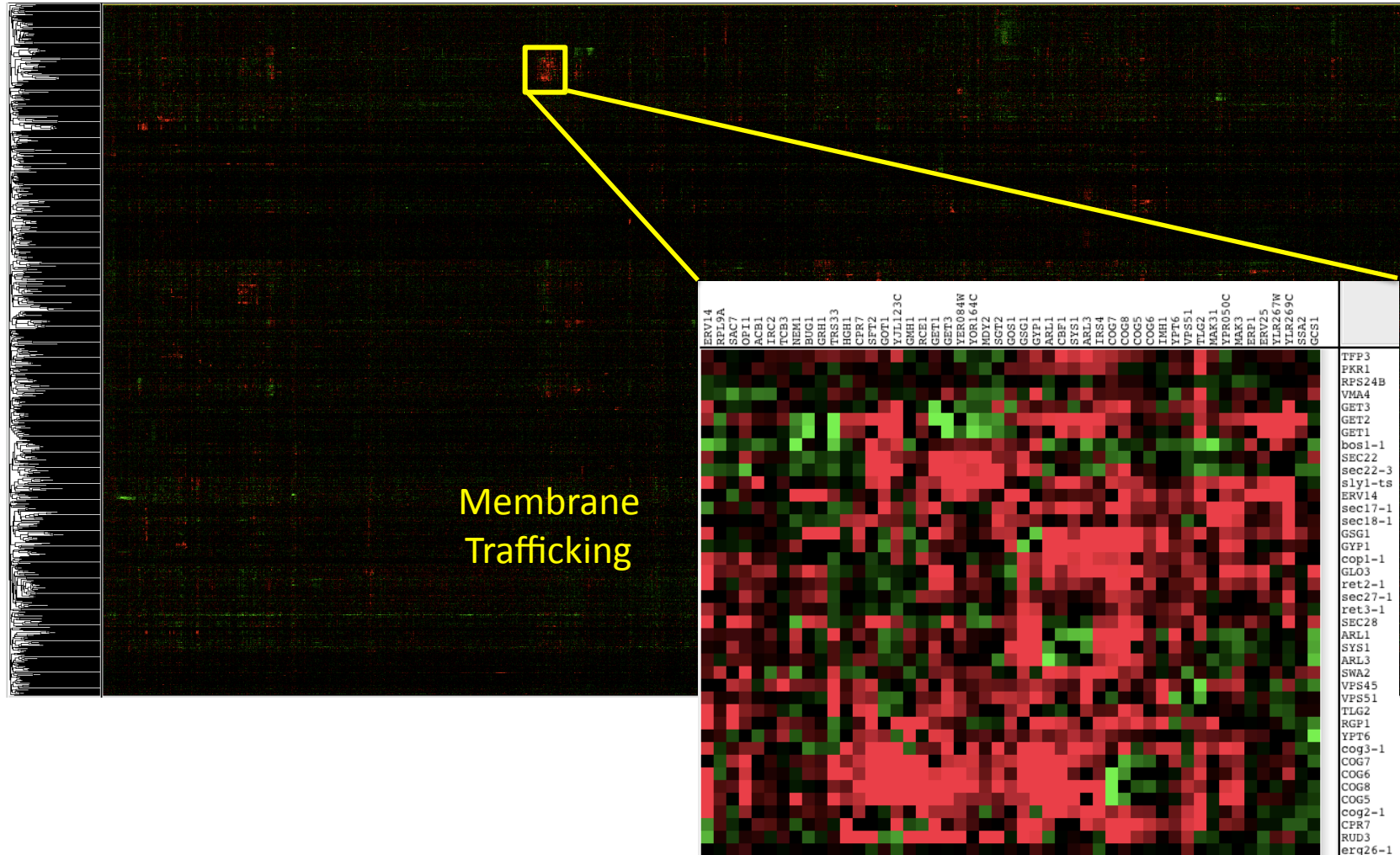
Olga Troyanskaya

Anastasia Baryshnikova



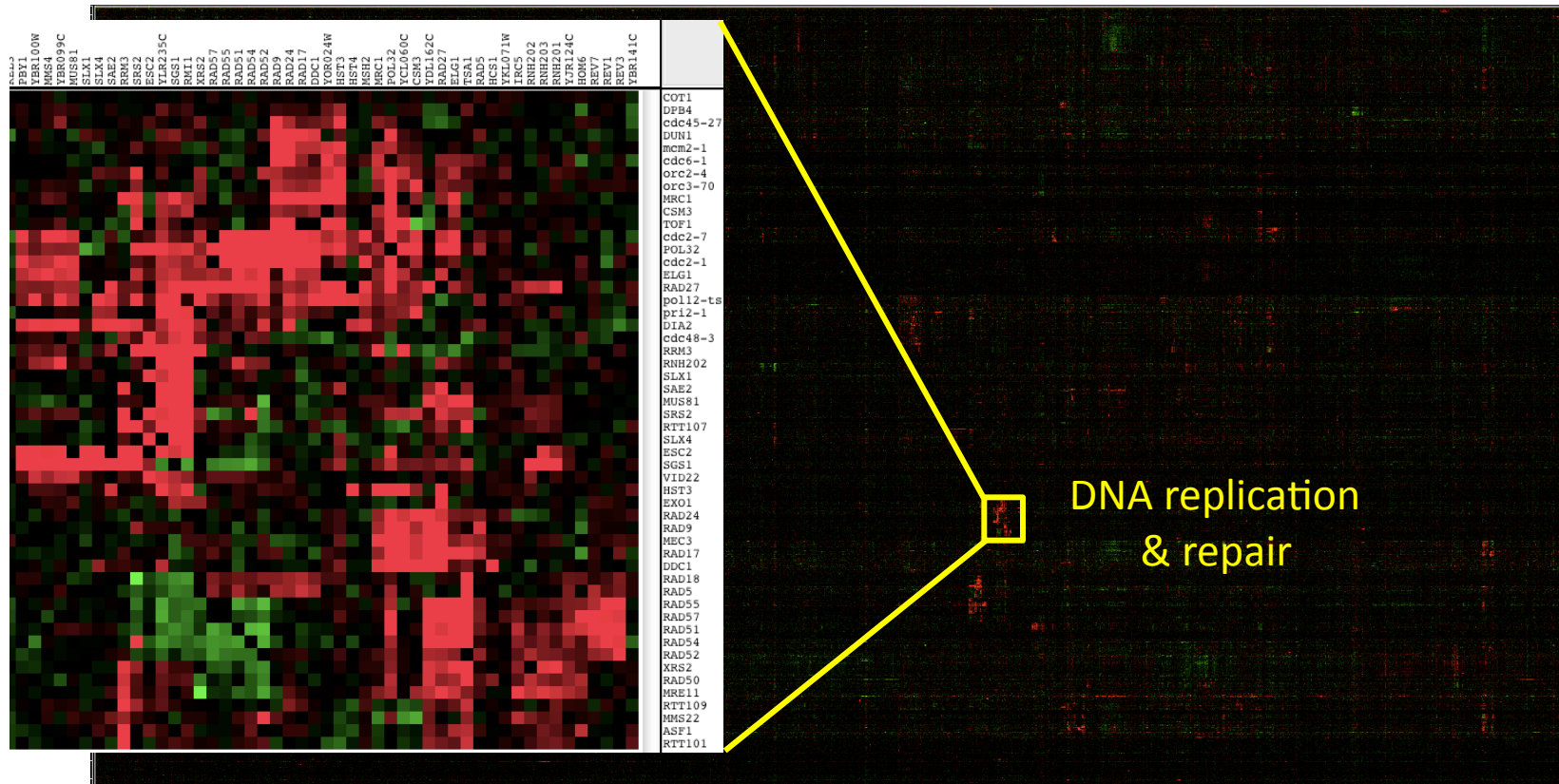
Full genome SGA matrix

994 x ~5000



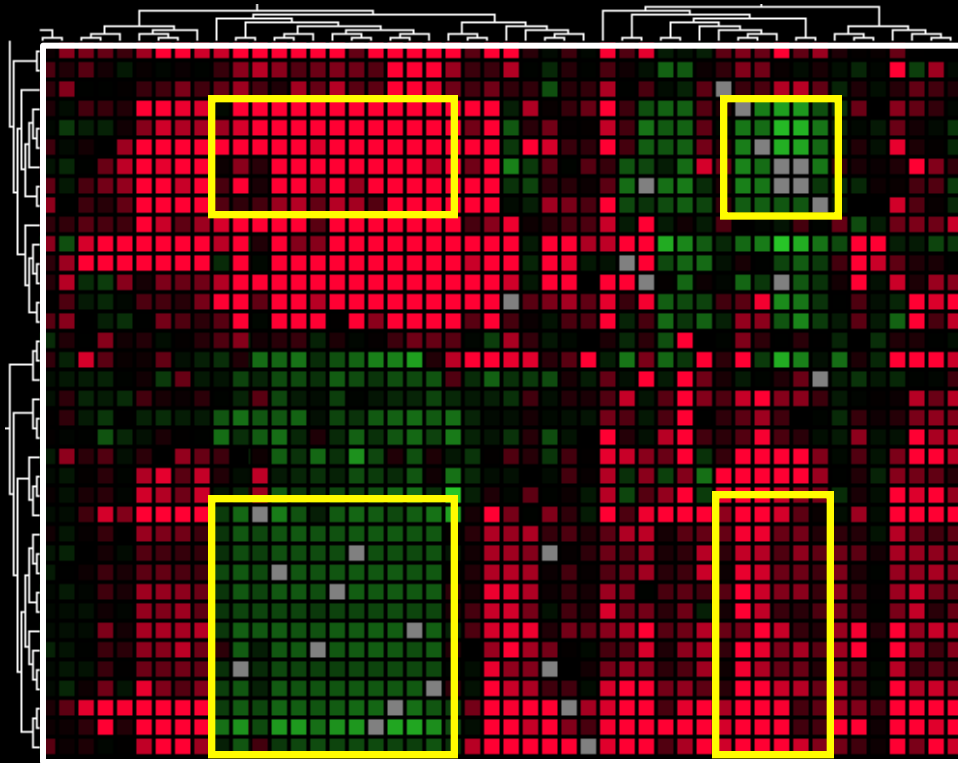
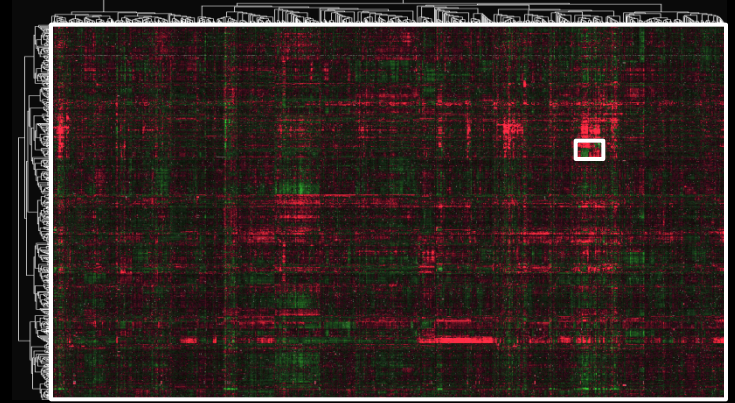
Full genome SGA matrix

994 x ~5000



SGA genetic interaction matrix

- aggravating interaction
- alleviating interaction



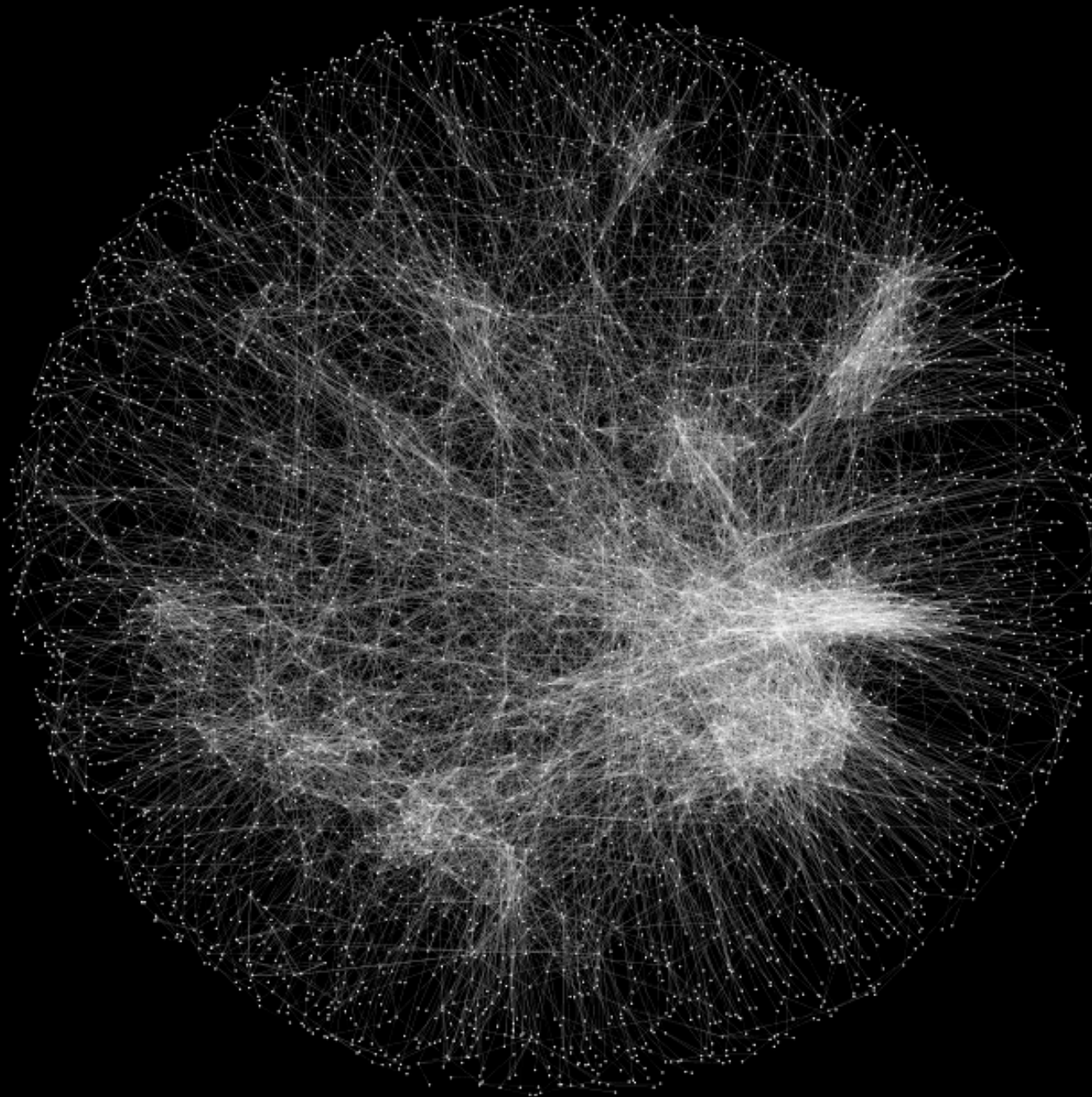
Retrograde-transport complex
(endosome-Golgi transport)

Vacuolar ATPase
(organelle acidification)

Vacuolar
ATPase

Retrograde
transport

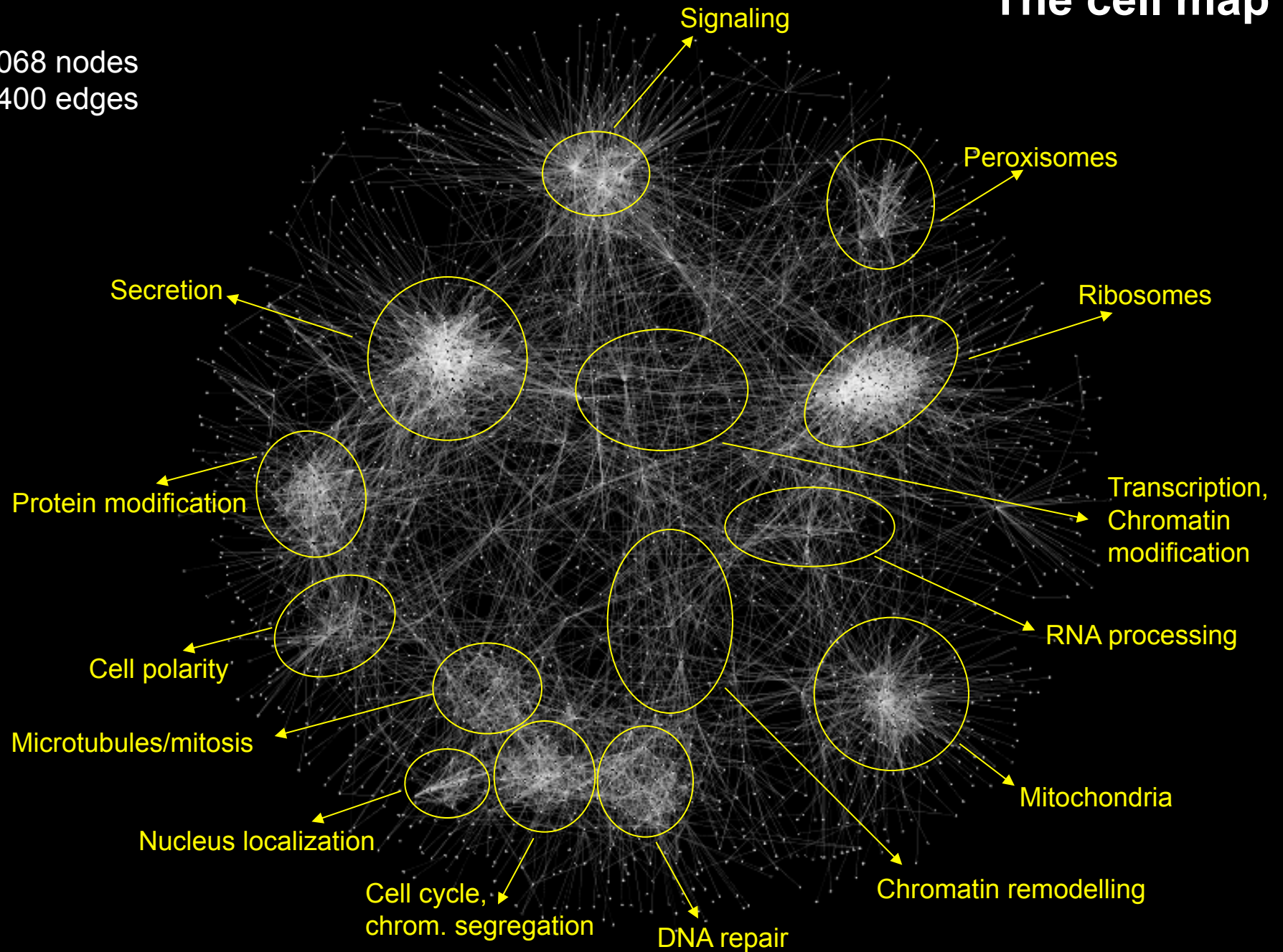
The cell map



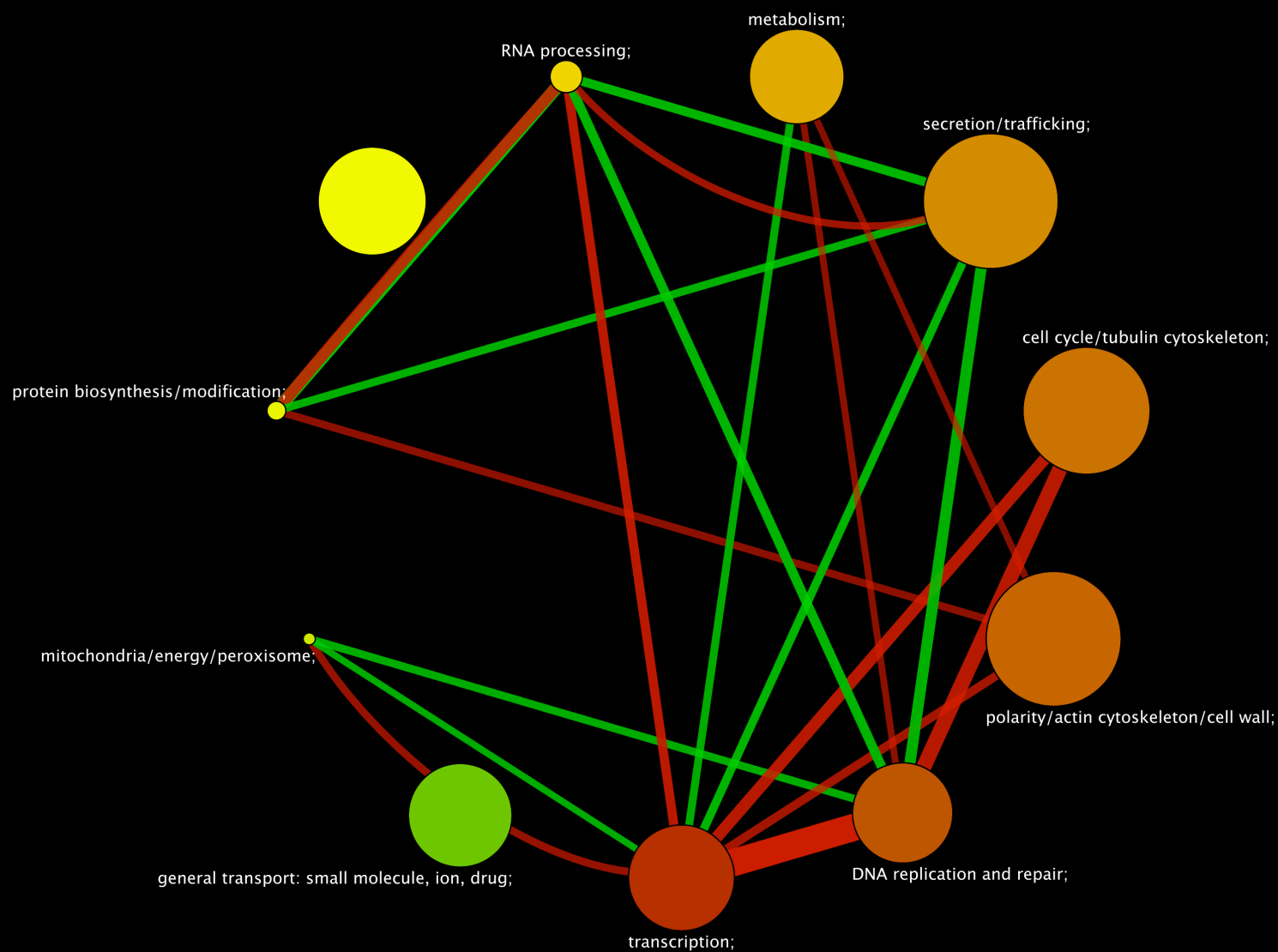
4000 nodes
13,754 edges

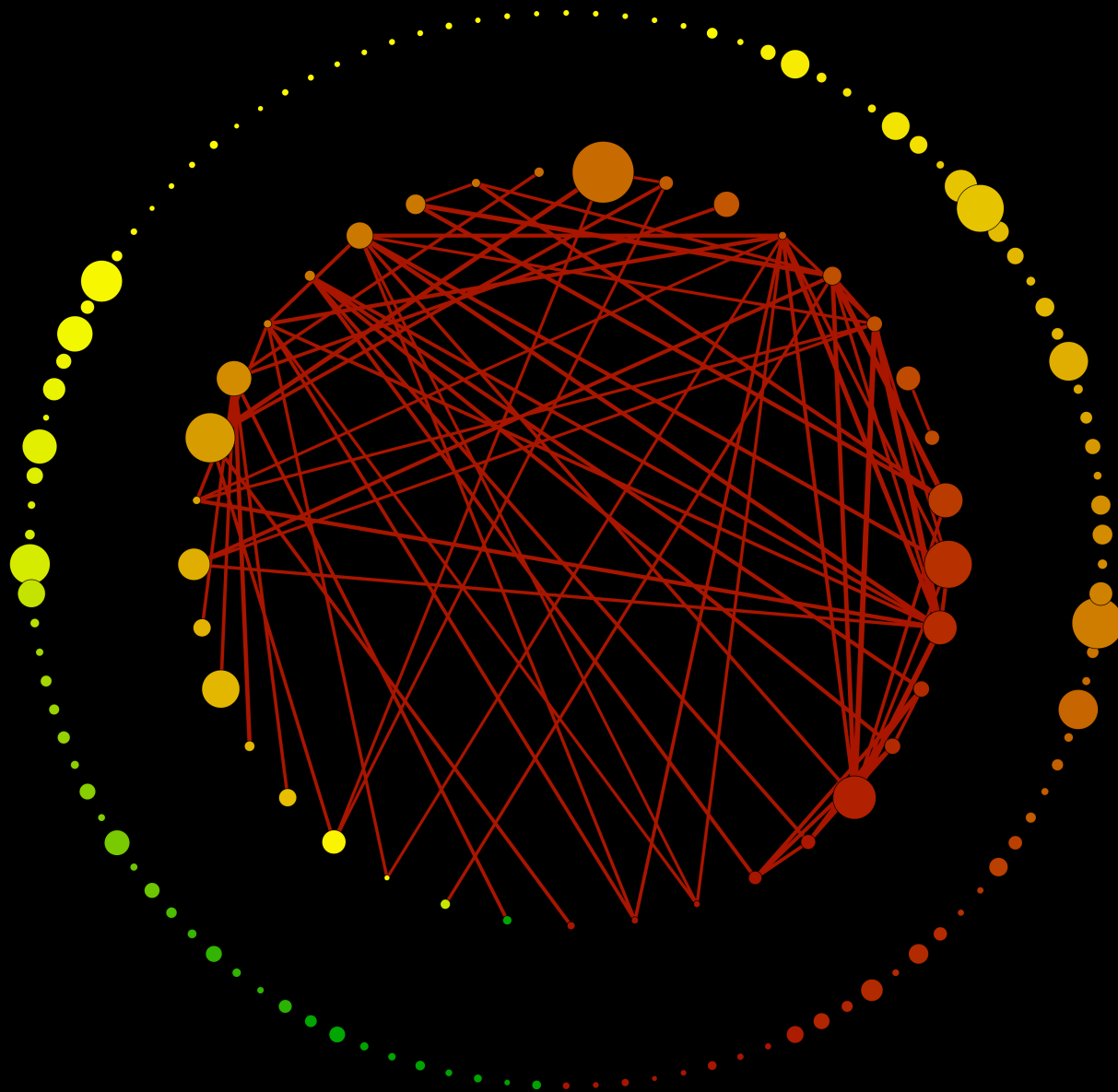
The cell map

2068 nodes
7400 edges



System Level Map of Yeast

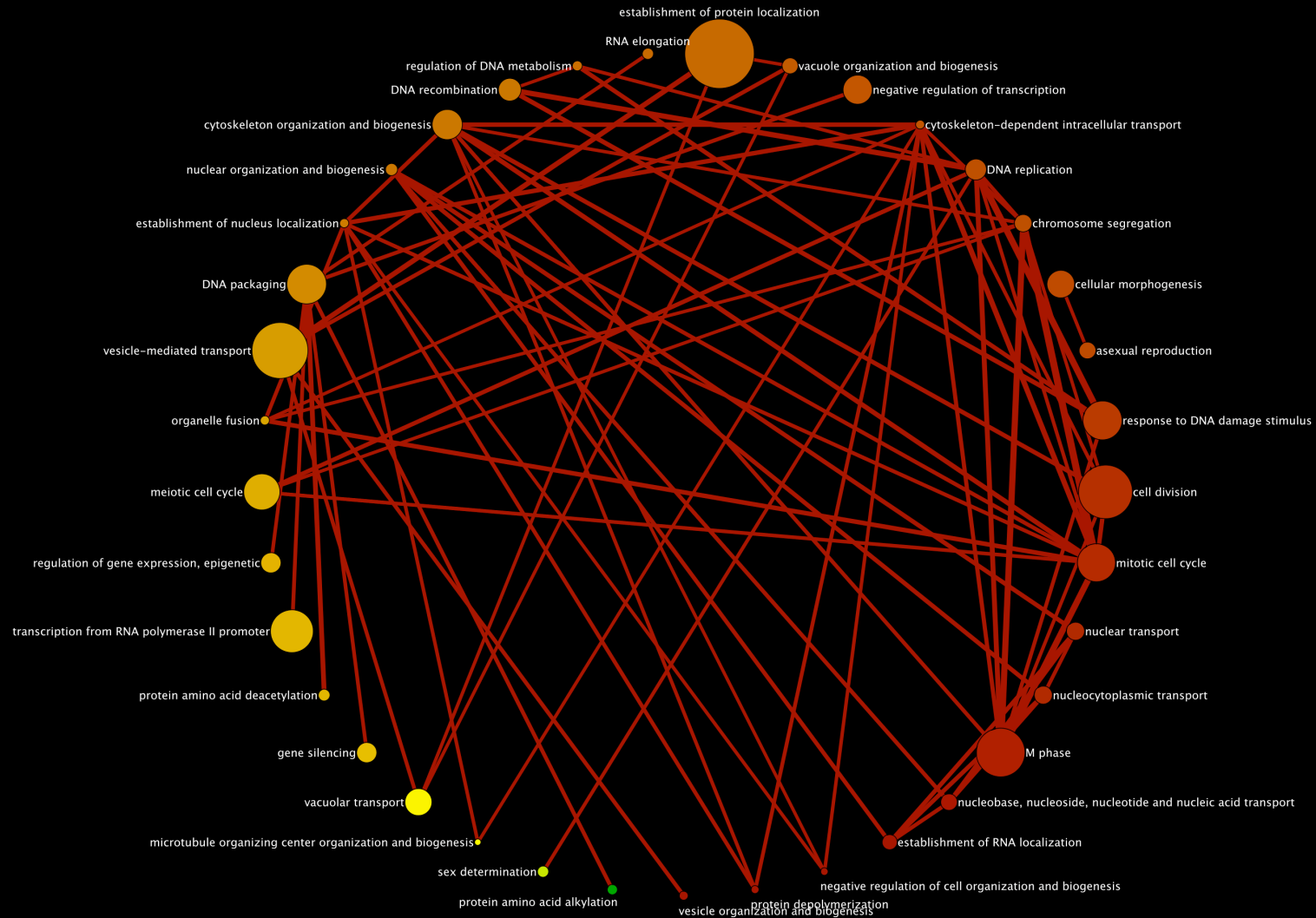




Systems Level Map

Edge Significance
Z-score > 10.0

Systems Level Map of Yeast



Genetic interactions with morphological profiling

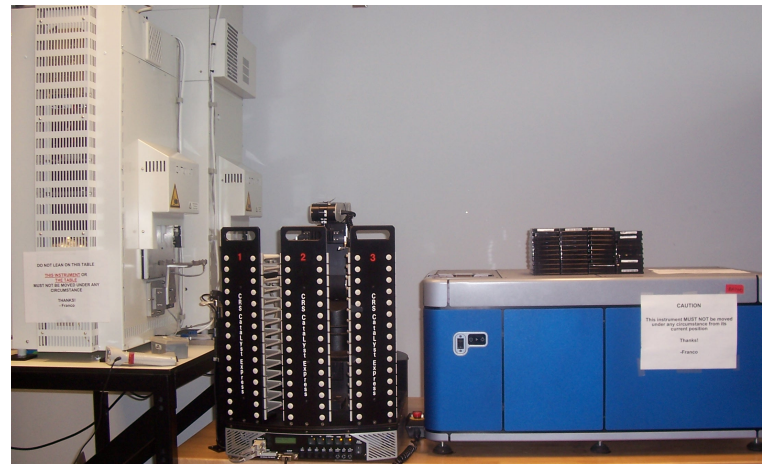
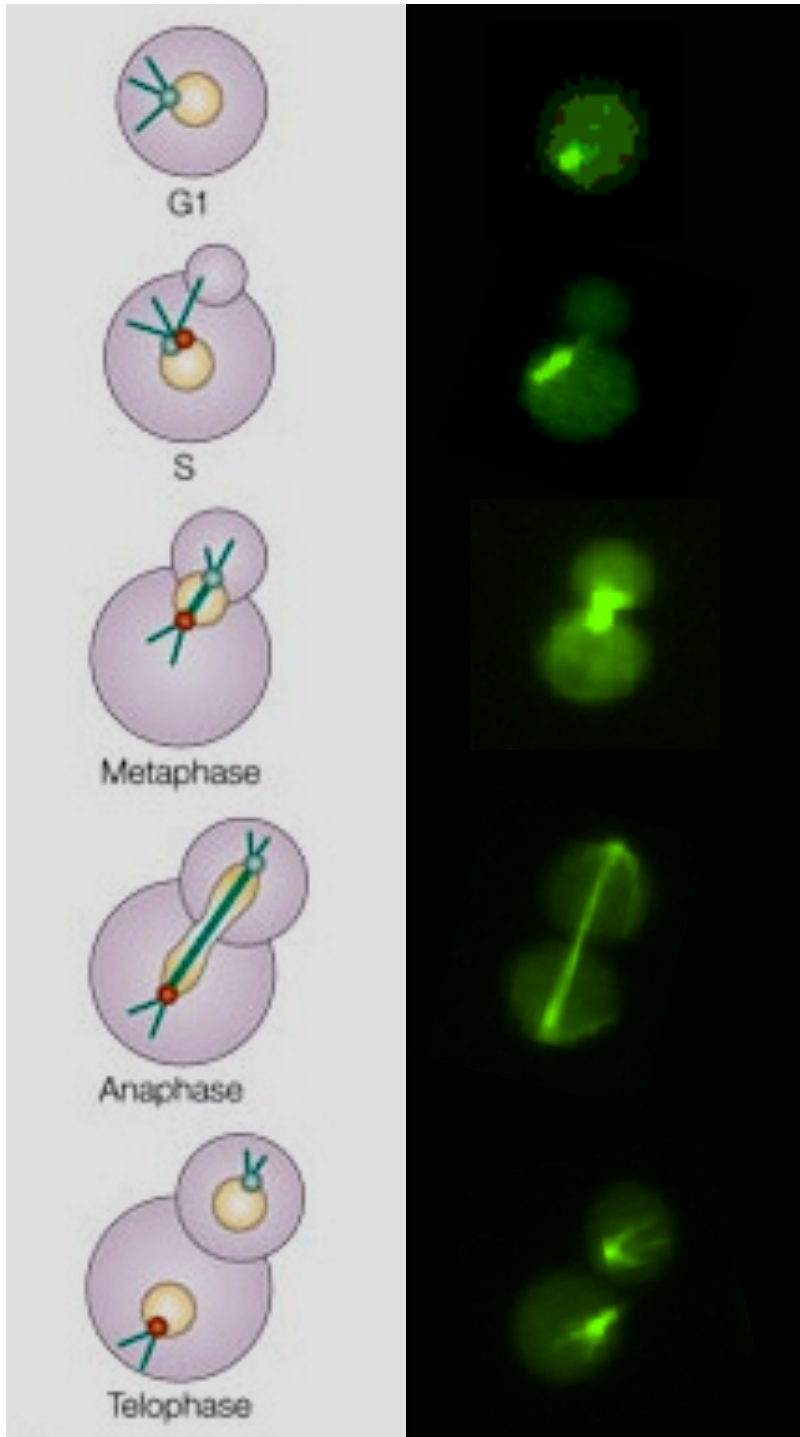
single mutants, bim1 & bni1 double mutants

40 parameters per cell

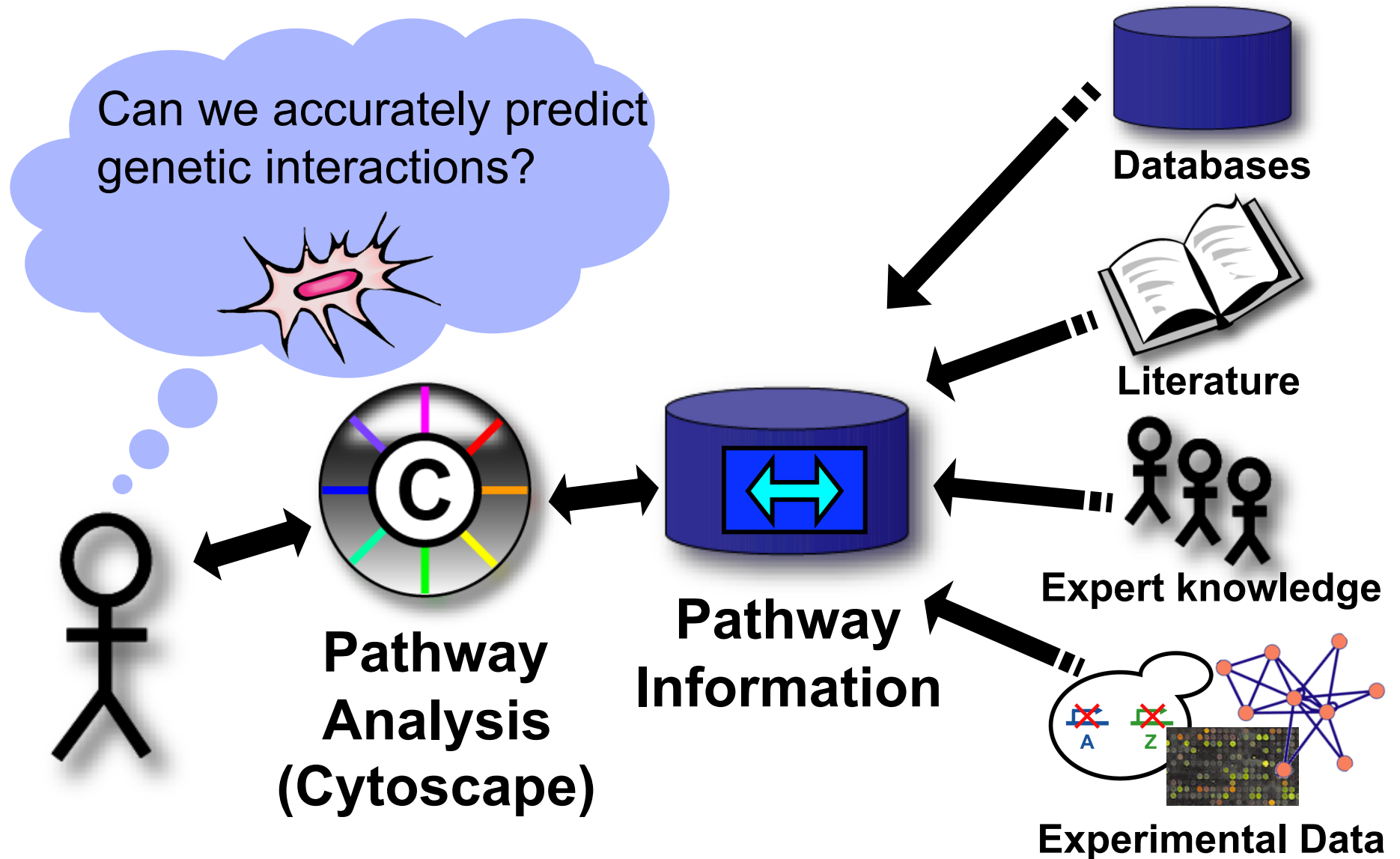
4,800-16,000 data points per mutant

~1,000 statistical parameters per mutant

- % unbudded, small, medium and large budded cells
- cell length, breadth, elliptical factor, area
- number of spindle pole bodies
- length, position, orientation of the spindle
- breadth of the budneck
- distance of the spindle from the budneck
- ...

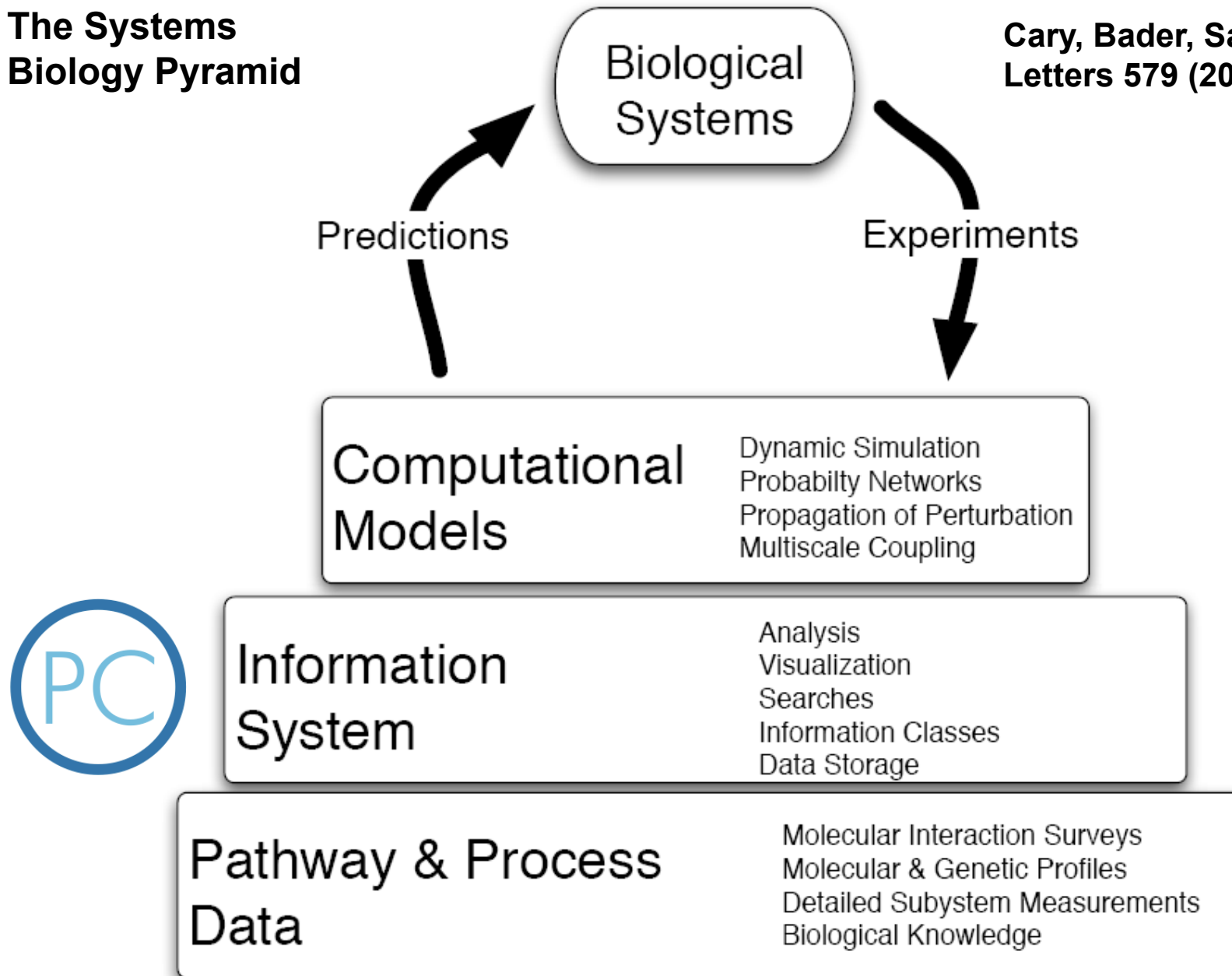


Cell map exploration and analysis



The Systems Biology Pyramid

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



http://pathguide.org

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](#).

~290 Pathway Databases!

Many new search options are available

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	PSI-MI
BindingDB - The Binding Database	Details	Free	
BioGRID - General Repository for Interaction Datasets	Details		PSI-MI
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	BioPAX
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		PSI-MI
Doodle - Database of oligomerization domains from lambda experiments	Details	Free	
DopaNet - DopaNet	Details	Free	
DRC - Database of Ribosomal Crosslinks	Details	Free	
DSM - Dynamic Signaling Maps	Details		
FIMM - Functional Molecular Immunology	Details	Free	
FusionDB - Prokaryote Gene Fusion Events	Details	Free	

Vuk Pavlovic

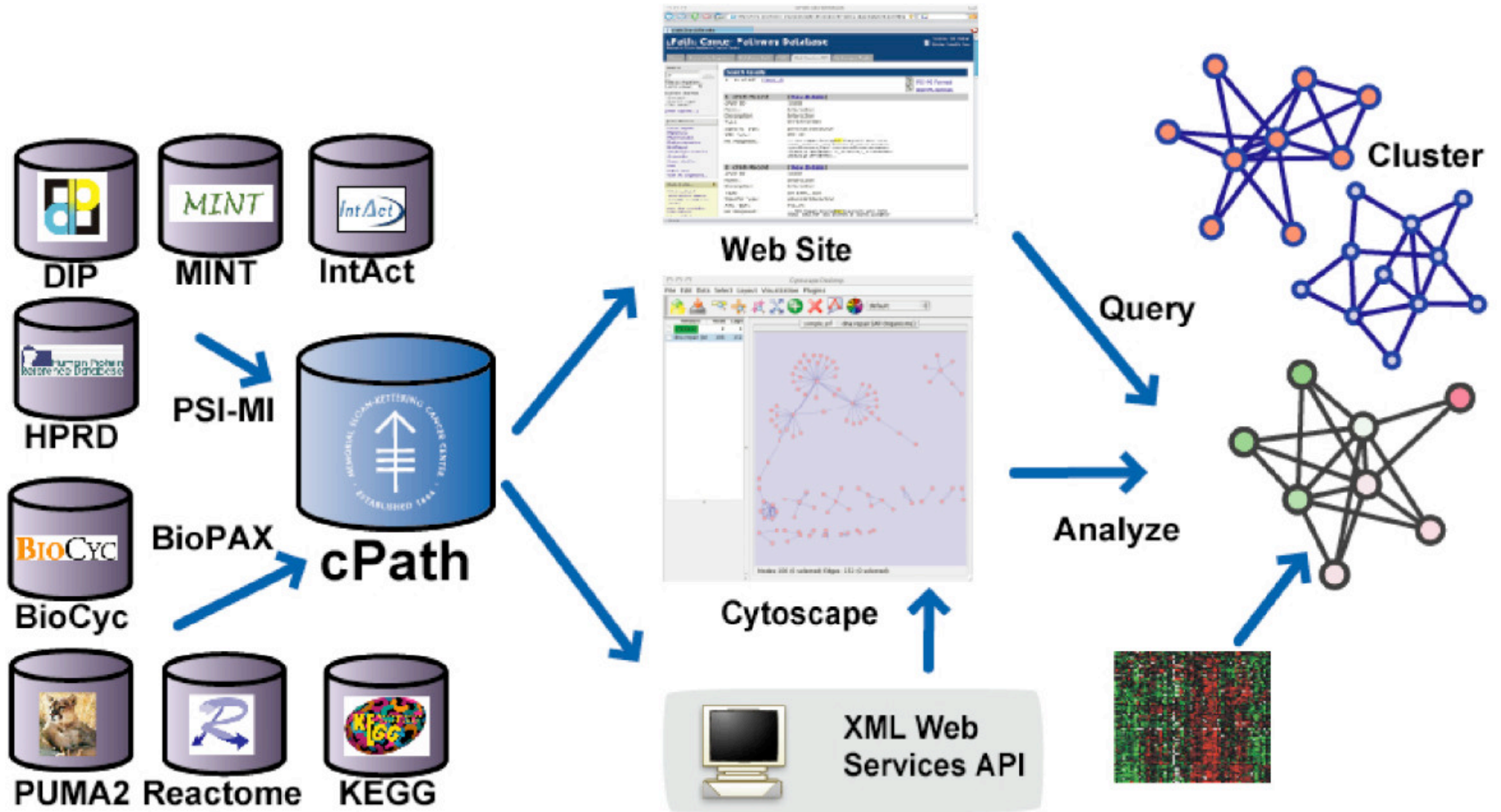
BioPAX Pathway Language

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

XML Snippet (OWL)

```
<bp:biochemicalReaction rdf:ID="biochemicalReaction37">
  <bp:DATA-SOURCE rdf:resource="#dataSource14"/>
  <bp:LEFT>
    <bp:physicalEntityParticipant rdf:ID="physicalEntityParticipant26">
      <bp:STOICHIOMETRIC-COEFFICIENT>1.0</bp:STOICHIOMETRIC-COEFFICIENT>
      <bp:PHYSICAL-ENTITY>
        <bp:smallMolecule rdf:ID="smallMolecule27">
          <bp:SHORT-NAME rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
            >a-D-glu-6-p</bp:SHORT-NAME>
          <bp:CHEMICAL-FORMULA rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
            >C6H13O9P</bp:CHEMICAL-FORMULA>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
            >&lt;FONT FACE="symbol">a&lt;/FONT>-D-glucose-6-phosphate</bp:SYNONYMS>
          <bp:XREF>
            <bp:unificationXref rdf:ID="unificationXref30">
              <bp:ID rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
                >C00668</bp:ID>
              <bp:DB rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
                >KEGG</bp:DB>
            </bp:unificationXref>
          </bp:XREF>
          <bp:XREF rdf:resource="#unificationXref29"/>
          <bp:MOLECULAR-WEIGHT>260.14</bp:MOLECULAR-WEIGHT>
          <bp:AVAILABILITY rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
            >see http://www.amaze.ulb.ac.be/</bp:AVAILABILITY>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
            >glucose-6-P</bp:SYNONYMS>
          <bp:DATA-SOURCE rdf:resource="#dataSource14"/>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
            >alpha-D-glucose-6-p</bp:SYNONYMS>
          <bp:STRUCTURE>
            <bp:chemicalStructure rdf:ID="chemicalStructure28">
              <bp:STRUCTURE-FORMAT>SMILES</bp:STRUCTURE-FORMAT>
              <bp:STRUCTURE-DATA>C(OP(=O)(O)O)[CH]1([CH](O)[CH](O)[CH](O)[CH](O)O1)</bp:STRUCTURE-DATA>
            </bp:chemicalStructure>
          </bp:STRUCTURE>
          <bp:NAME>alpha-D-glucose 6-phosphate</bp:NAME>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
            >alpha-D-glucose-6-phosphate</bp:SYNONYMS>
          <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
            >D-glucose-6-P</bp:SYNONYMS>
          <bp:DATA-SOURCE rdf:resource="#KB_439584_Individual_47"/>
        </bp:smallMolecule>
      </bp:PHYSICAL-ENTITY>
      <bp:CELLULAR-LOCATION rdf:resource="#openControlledVocabulary15"/>
    </bp:physicalEntityParticipant>
  </bp:LEFT>
  <bp:DELTA-G rdf:datatype="http://www.w3.org/2001/XMLSchema#double"
    >0.4</bp:DELTA-G>
  <bp:SYNONYMS rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
    >alpha-D-Glucose 6-phosphate &lt;=> beta-D-Fructose 6-phosphate </bp:SYNONYMS>
  <bp:RIGHT>
    <bp:physicalEntityParticipant rdf:ID="physicalEntityParticipant38">
      <bp:CELLULAR-LOCATION rdf:resource="#openControlledVocabulary15"/>
      <bp:PHYSICAL-ENTITY>
        <bp:smallMolecule rdf:ID="smallMolecule39">
```

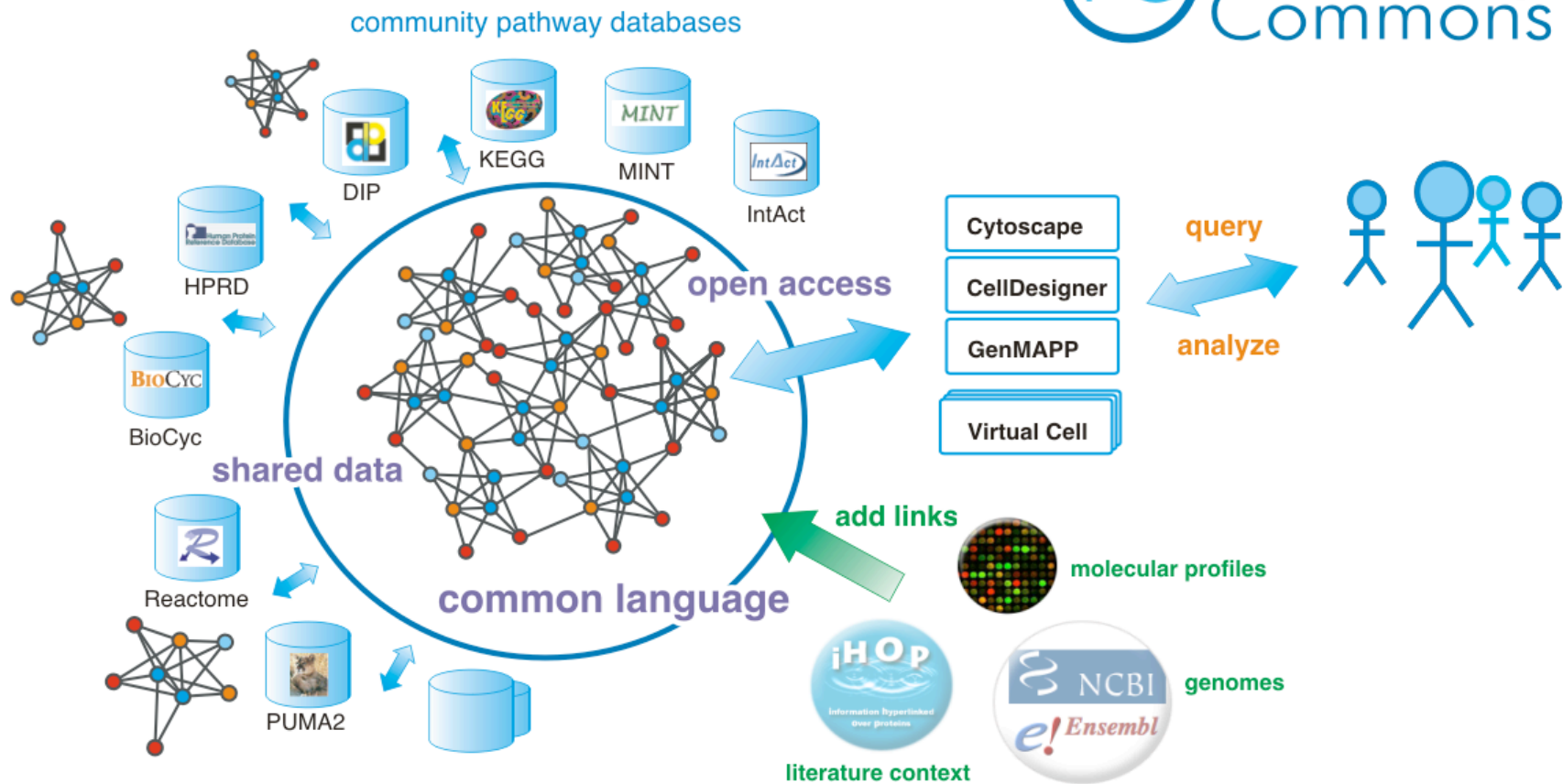
cPath Pathway Database Software



Collect → **Browse / Query** → **Analyze**

Pathway Commons: A Public Library

<http://pathwaycommons.org>



- Books: Pathways
- Lingua Franca: BioPAX

- Index: cPath pathway database software
- Translators: translators to BioPAX

<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

Cytoscape Desktop (New Session)

File Edit View Select Layout Plugins Help

Sample2 Search: yal003w

CytoPanel 1

Network Editor

Network	Nodes	Edges
galFiltered.sif	331(19)	362(35)

CytoPanel 2

Node Attribute Browser (galFiltered.sif)

ID	gal1RGexp	gal1RGsig	gal4RGexp	gal4RGsig	gal80Rexp	gal80Rsig
YGL008C	-0.352	1.0007E-5	-0.282	7.1366E-4	-0.573	1.2622E-5
YCL067C	0.169	0.0012873	-0.085	0.11481	0.301	0.0027E-5
YNL145W	-0.764	3.148E-11	-0.098	0.05338	-1.237	1.1916E-5
YMP043W	-0.183	0.0035372	-0.654	4.2514E-6	0.457	2.4112E-5

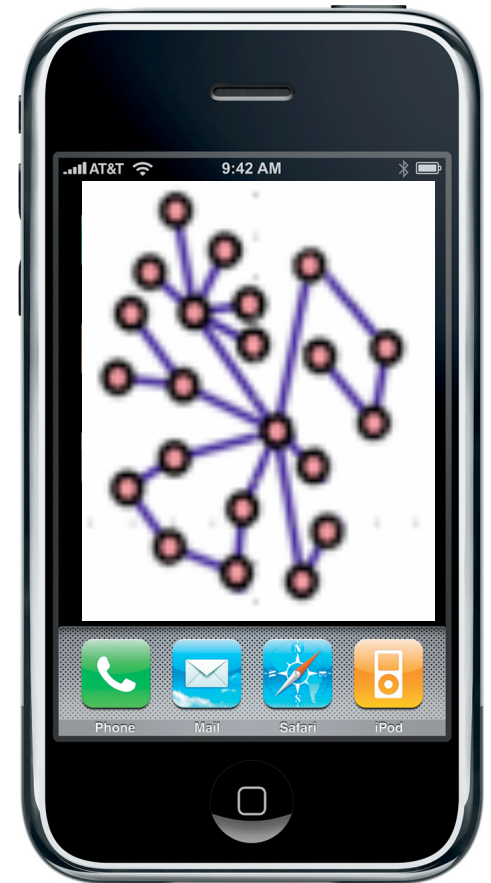
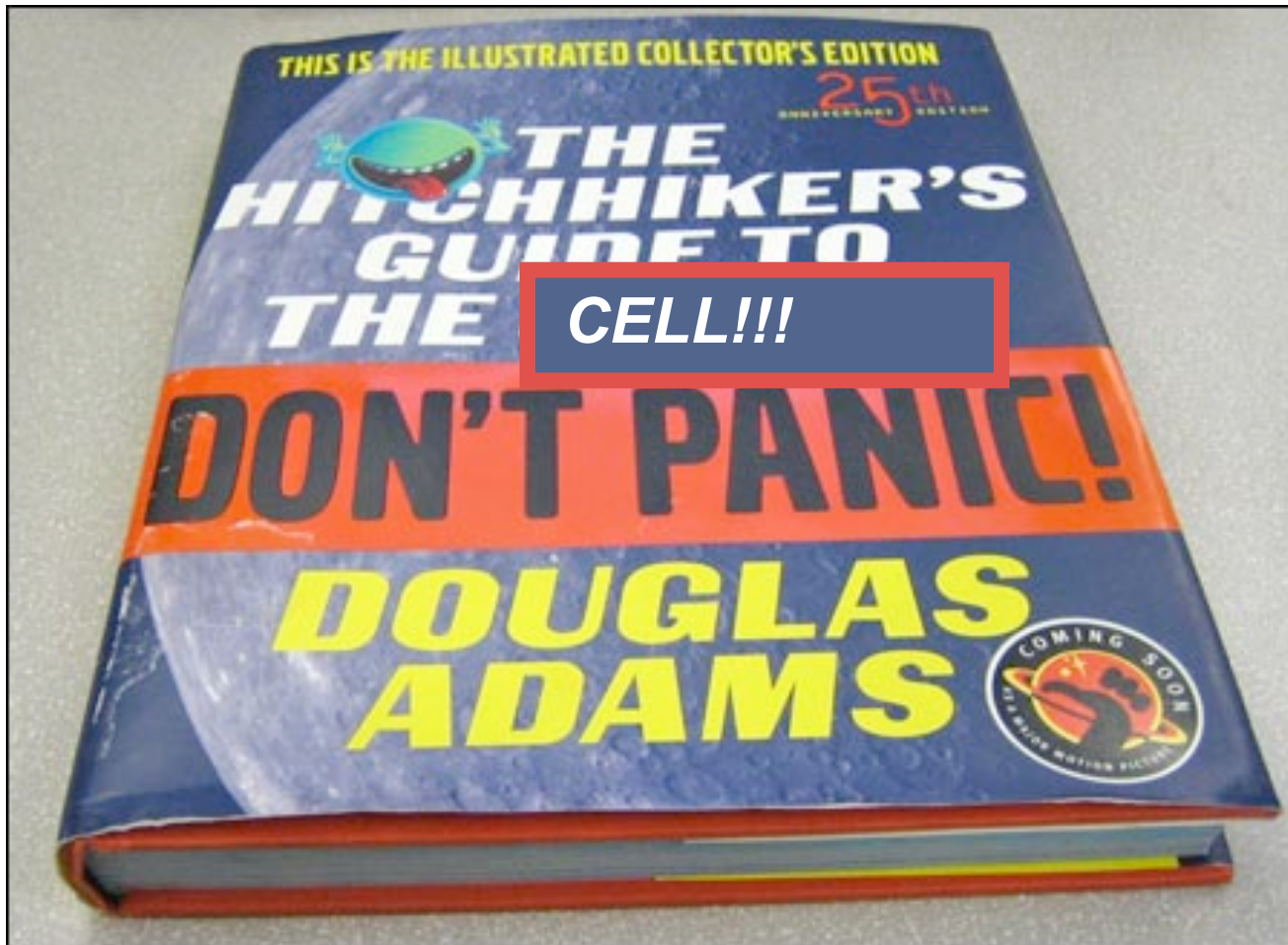
Welcome to Cytoscape 2.4.0-b1

Right-click + drag to ZOOM Middle-click + drag to PAN

Challenges

- Data: Author entry systems
 - From individual publications
 - For pathways (review)
 - Curator tools (advanced)
- Semantic integration (Identifier resolution)
- Visualization
 - Pathway diagrams (SBGN)
 - Automated layout
- Algorithms for compound graphs
- Linking discrete and dynamic representations
 - Including use by modelers

Where we want to be with
cellular visualization...



Total Video Converter
<http://effectmatrix.com>

Acknowledgements

Genetic Interactions

Charlie Boone (UofT)

Anastasia Baryshnikova

Michael Costanzo

Huiming Ding

Judice L. Y. Koh

Kiana Toufighi

Brenda Andrews (UofT)

Chad Myers (UMinn)

**Olga Troyanskaya
(Princeton)**

Funding

CIHR, NSERC, NIH

Genome Canada

Canada Foundation
for Innovation/ORF

Bader Lab

MP

Anastasija Baryshnikova

Ben Elliott

Iain Wallace

Laetitia Morrison

G2N

Chris Tan

David Gfeller

Moyez Dharsee

Shirley Hui

ACM

Daniele Merico

Ruth Isserlin

Vuk Pavlovic

Pathway Commons

Chris Sander

Ethan Cerami

Ben Gross

Emek Demir

Robert Hoffmann

Cytoscape

Trey Ideker (UCSD)

Ryan Kelley, Kei Ono, Mike Smoot, Peng
Liang Wang (Nerius Landys, Chris
Workman, Mark Anderson, Nada Amin,
Owen Ozier, Jonathan Wang)

Lee Hood (ISB)

Sarah Killcoyne (Iliana Avila-Campillo,
Rowan Christmas, Andrew Markiel, Larissa
Kamenkovich, Paul Shannon)

Benno Schwikowski (Pasteur)

Melissa Cline, Tero Aittokallio

Chris Sander (MSKCC)

Ethan Cerami, Ben Gross (Robert Sheridan)

Annette Adler (Agilent)

Allan Kuchinsky, Mike Creech (Aditya
Vailaya)

Bruce Conklin (UCSF)

Alex Pico, Kristina Hanspers

<http://baderlab.org>

