Generating Tailored Solutions from Big Data:
The Future of Health and Environmental Data Interrogation

In Silico Modeling of Biological Networks

Nathan Price

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

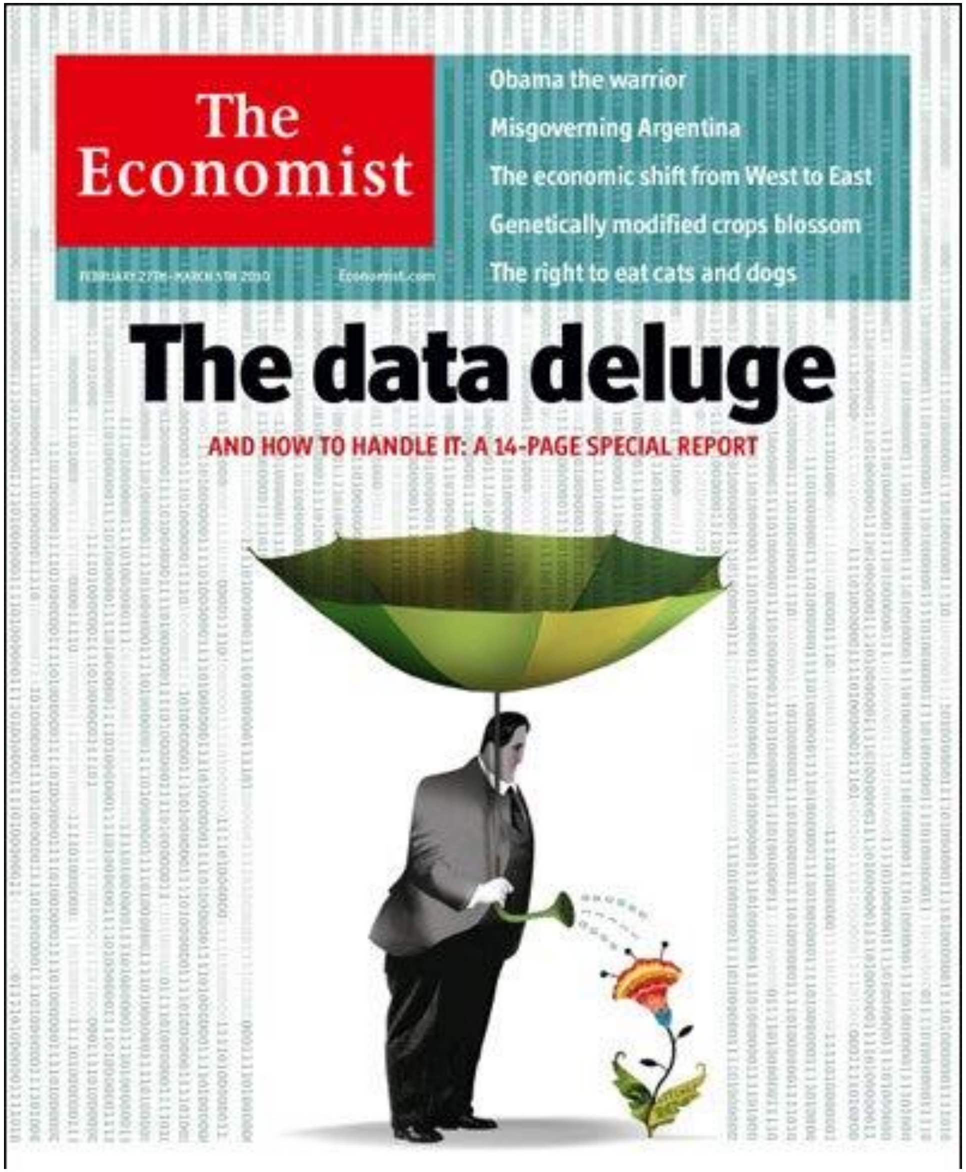
Disclosure

- Dr. Price has a significant financial interest in a commercial entity which partially funds, and which may license discoveries resulting from, the Hundred Person Wellness Project (to be described).

Extracting knowledge from data

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

We are NOT drowning in data in biology



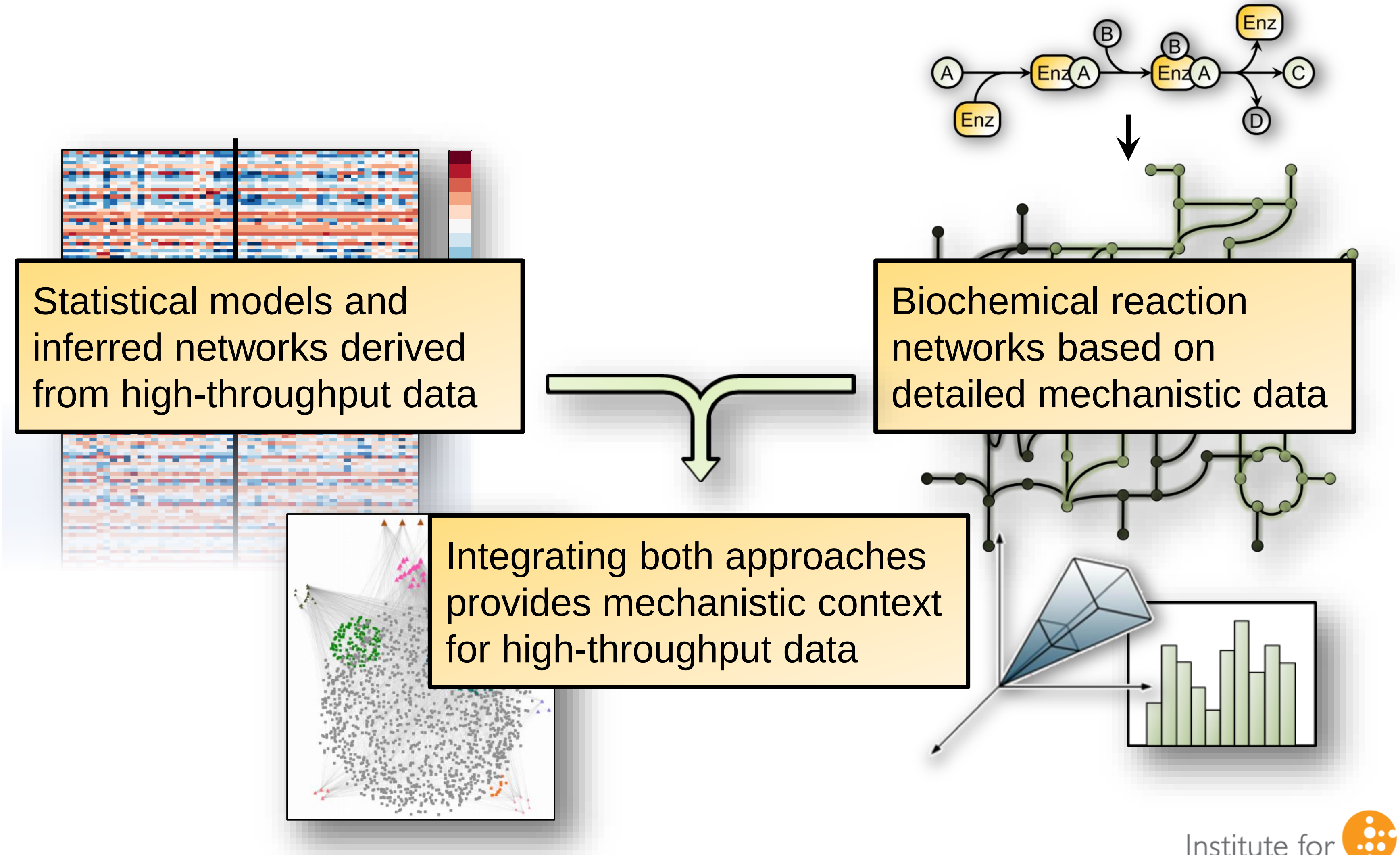
Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Dealing with uncertainty is an essential aspect of computational biology

Network models as complex, rigorously structured hypotheses

Institute for
Systems Biology
Revolutionizing Science. Enhancing Life.

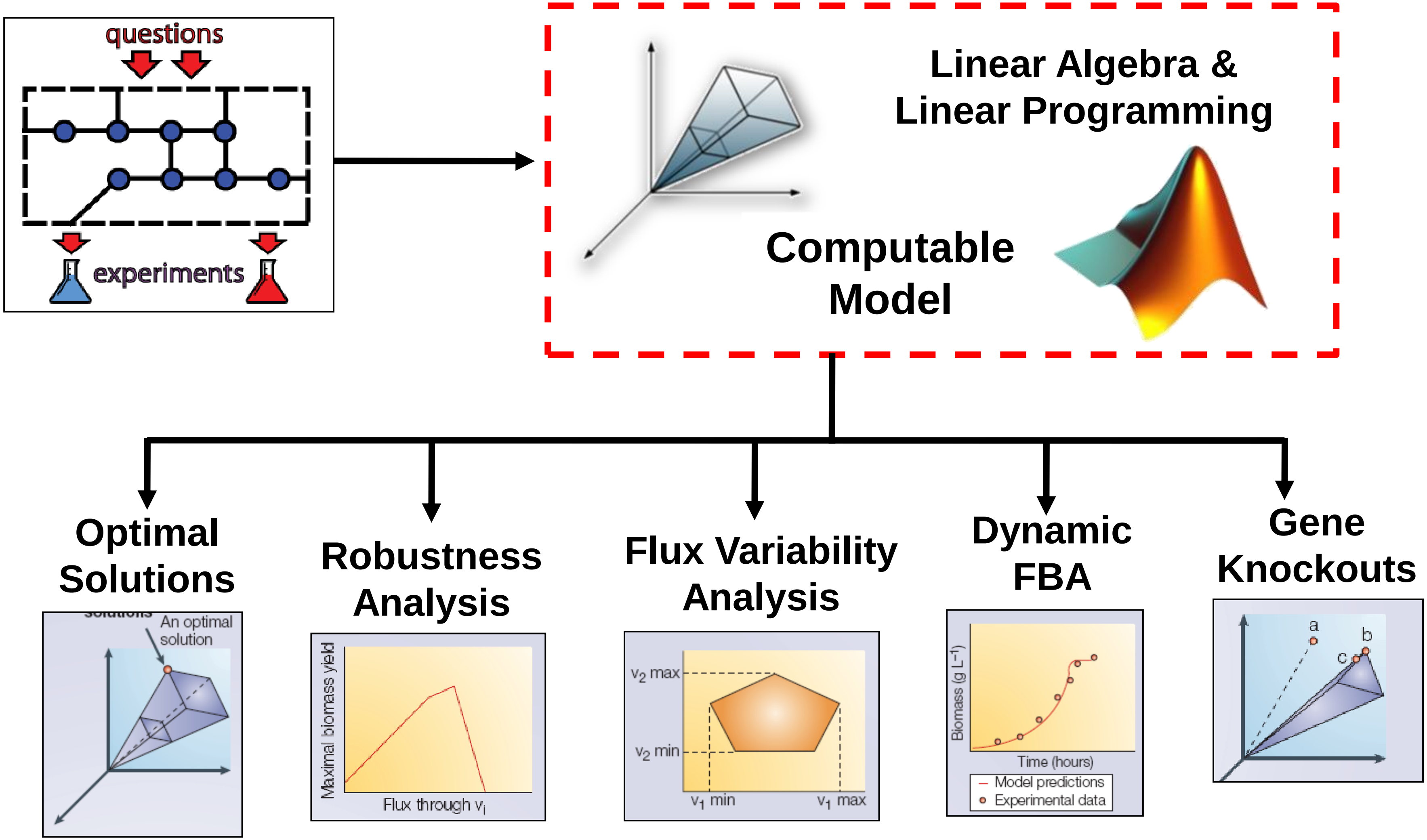
Statistical and mechanistic models provide context for data interpretation



Price, N.D., and Shmulevich, I., *Current Opinion in Biotechnology*, 2007

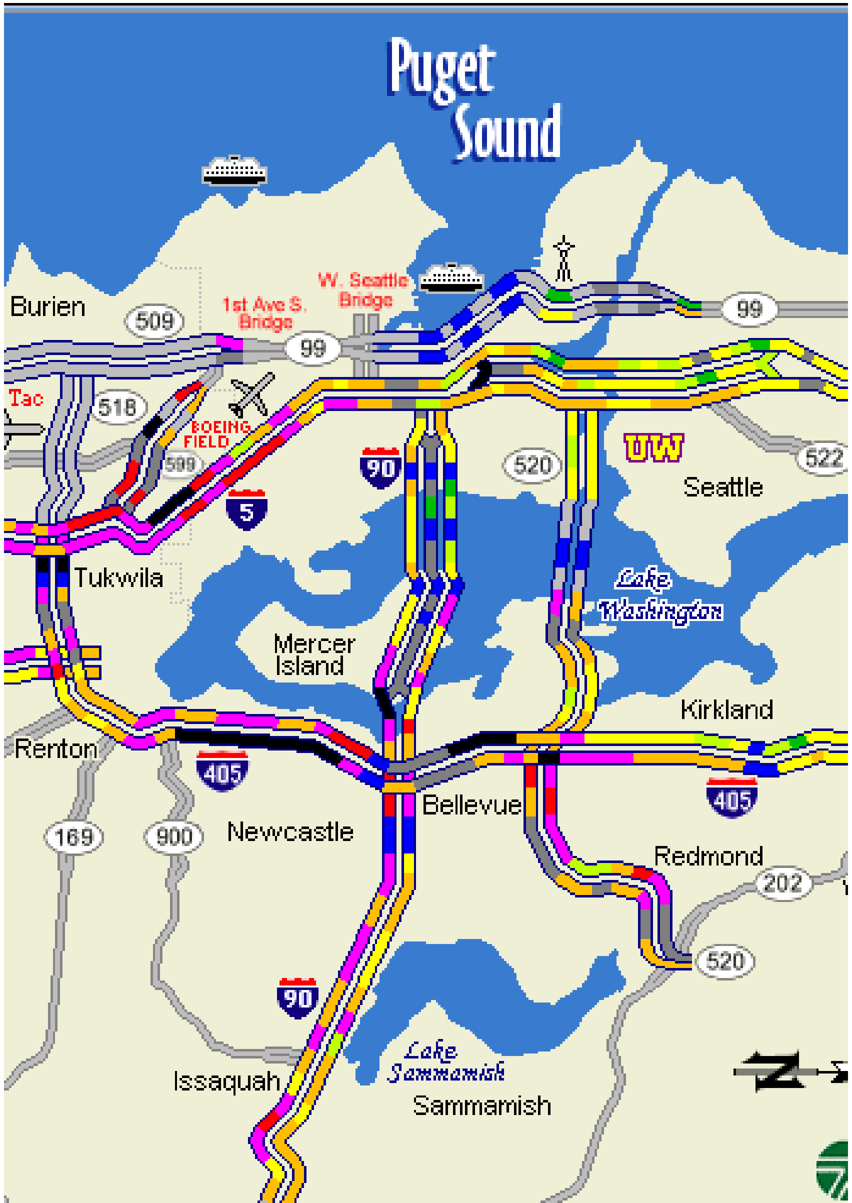
Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

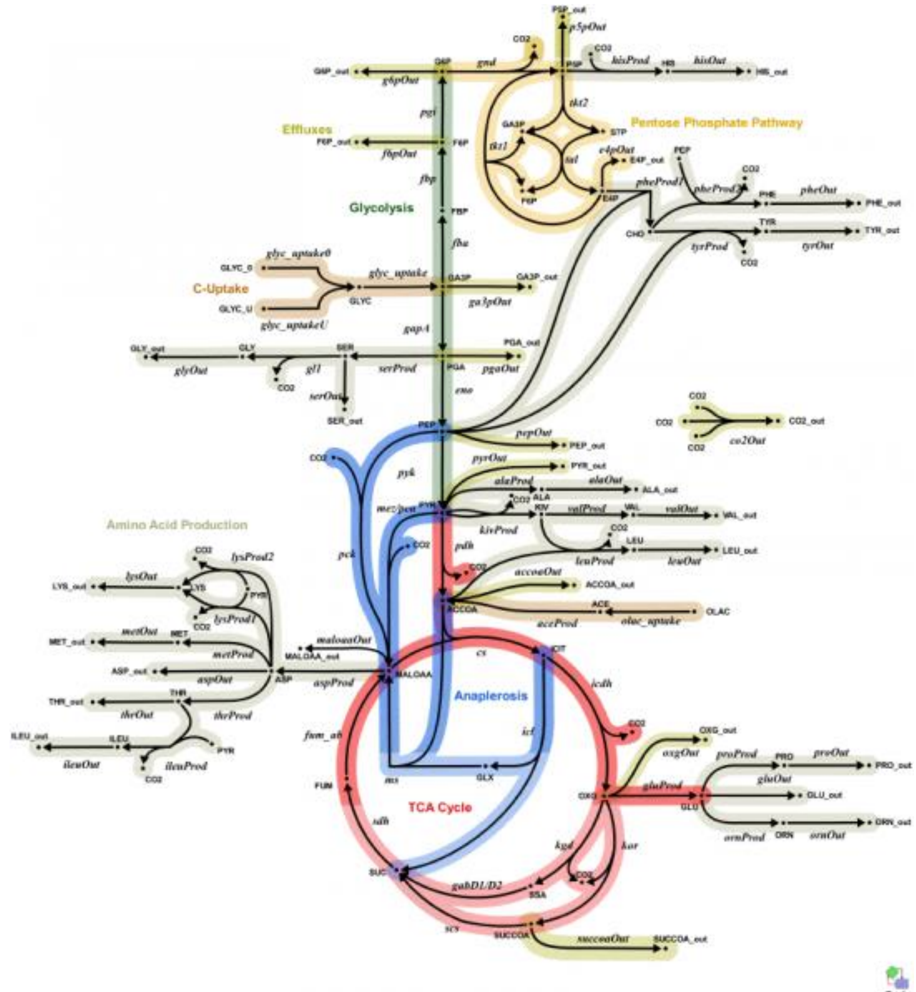
Model simulations allow for phenotype prediction



Becker, S. A. F. et al. 2007. Nature Protocols 23: 727-738.
Price, N.D. et al.. 2004. Nature Reviews.2: 886

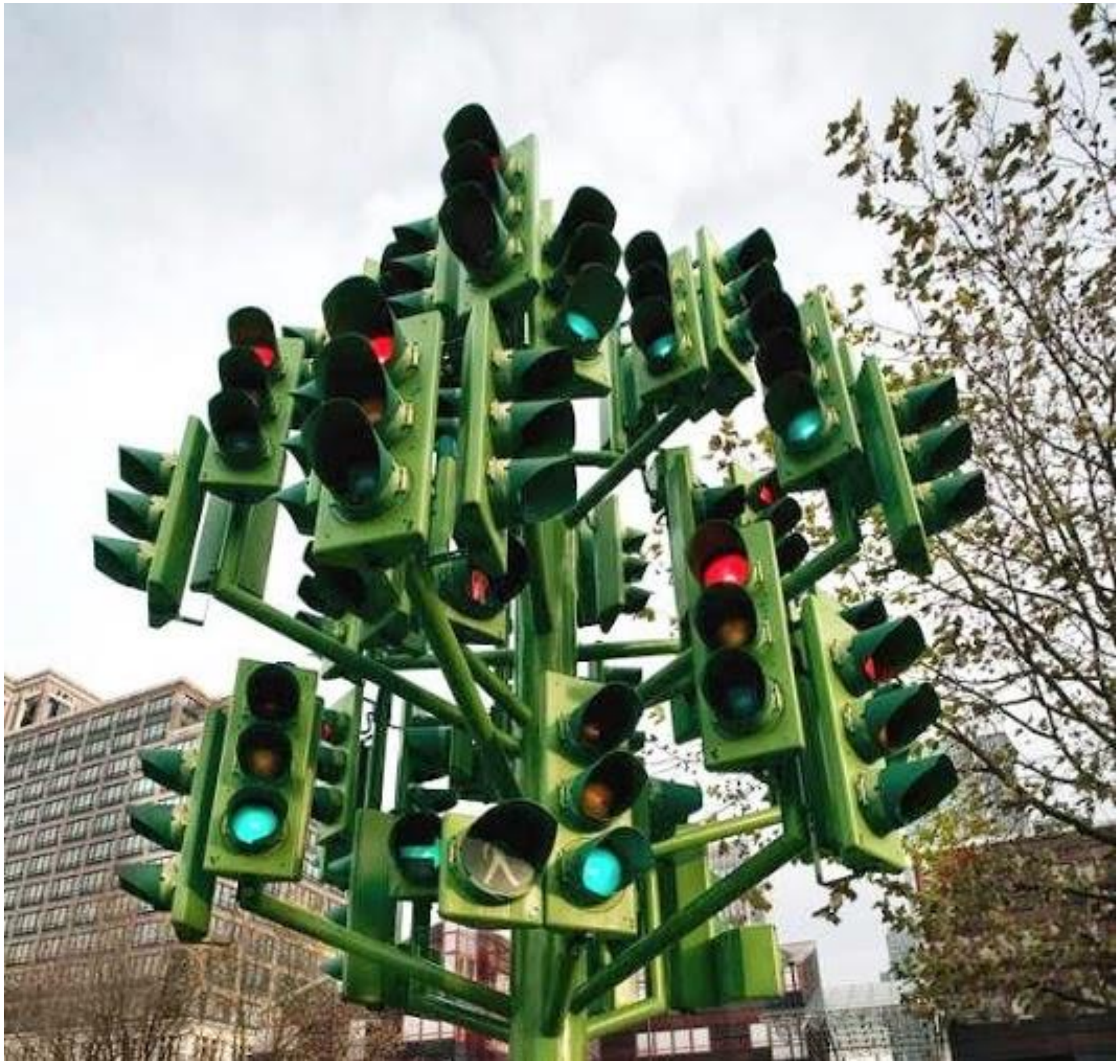
Metabolic Flux: 'Biological Traffic Flow'





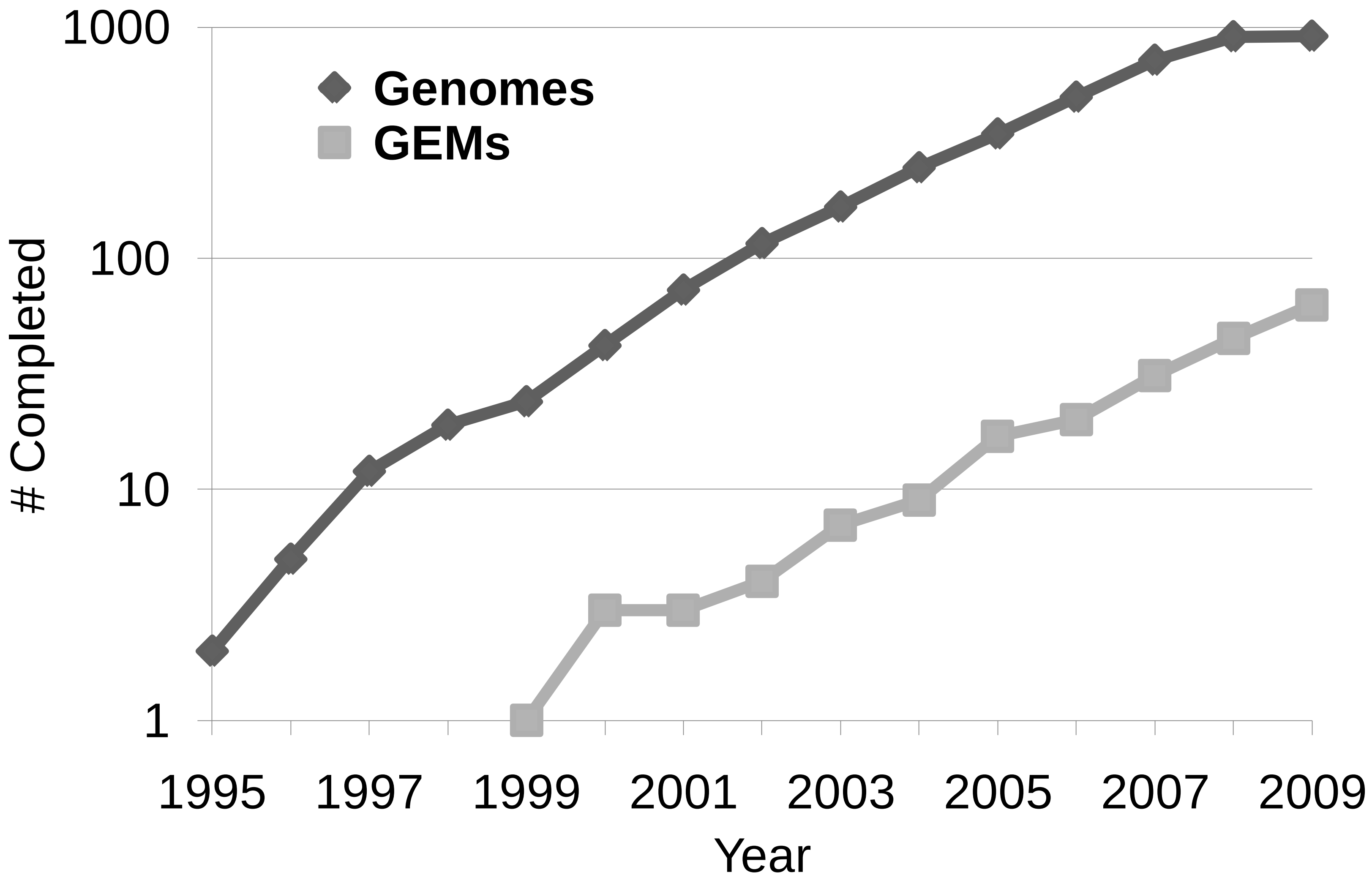
Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Complex array of regulatory control for metabolic traffic



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Need for automated network reconstruction

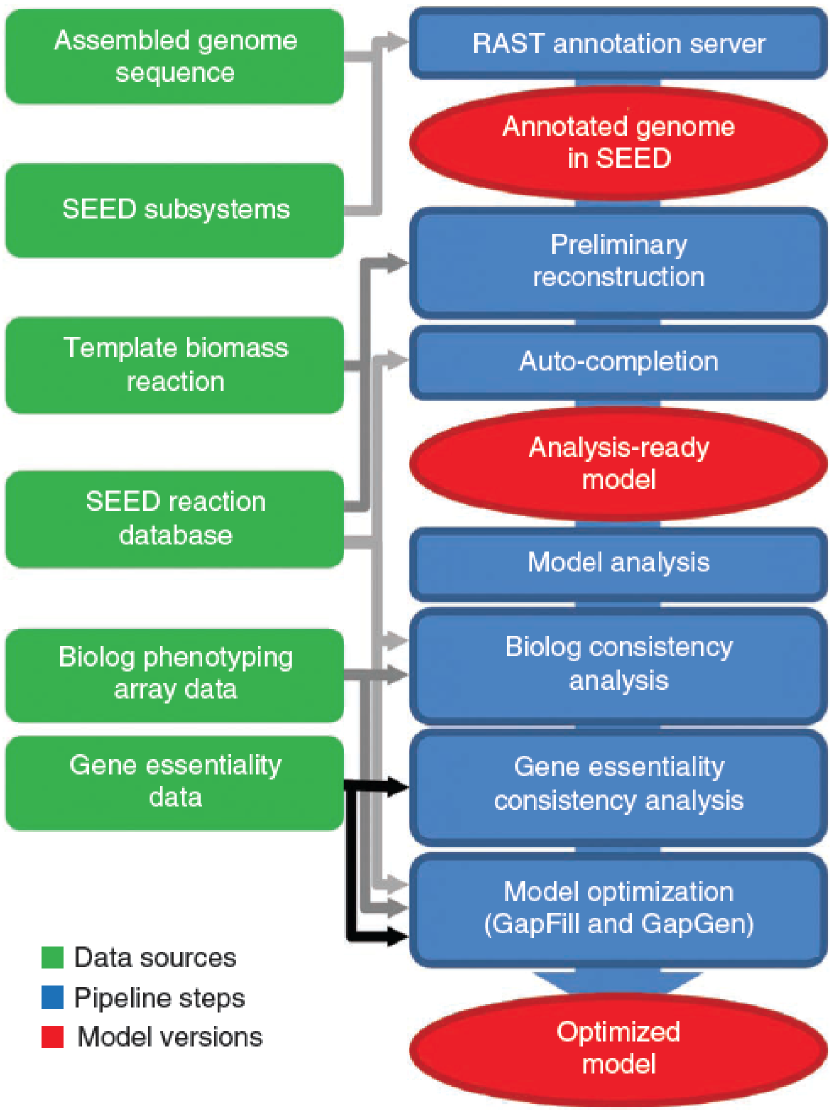


C Milne, JA Eddy, PJ Kim, ND Price, *Biotechnology Journal*, 2009

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Automated reconstruction of metabolic networks

- Automated reconstruction of *computable* metabolic network models
- Demonstrated on 130 genomes – now done for ~3000
- Provide advanced starting point for virtually any organism
- Accuracy from genomics: 65%
- With biolog and optimization: 87%



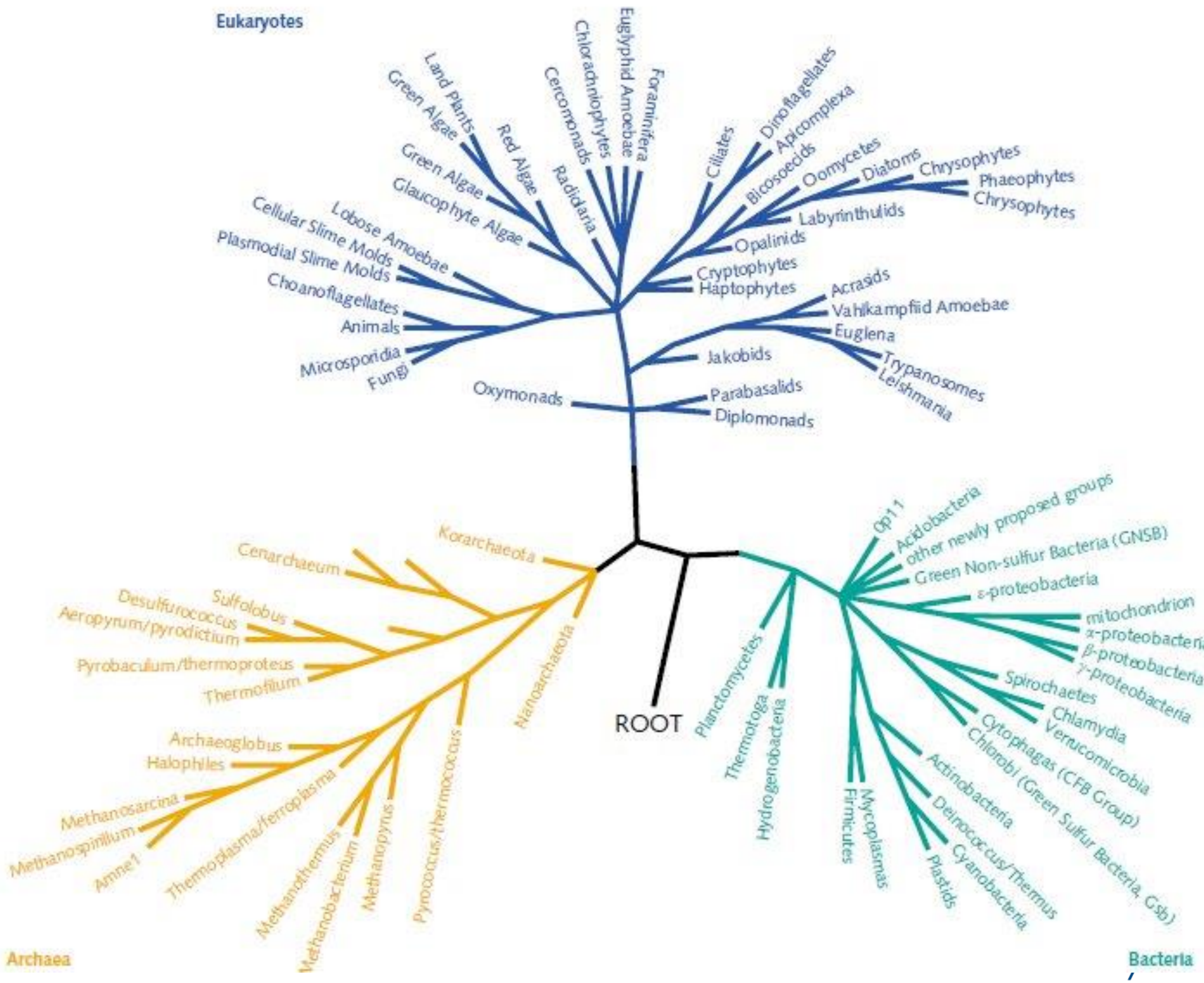
- Henry, C. DeJongh, M, Best, AA, Frybarger, PM, and Stevens, RL, *Nature Biotechnology*, 2010

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Enormous scope of microbiomes motivates scalable approaches

earth
microbiome project

Metabolic network expansion around key species in phylogenic tree



stitute for
ns Biology
Revolutionizing Science. Enhancing Life.

Microbes are Critical Components of the Global Carbon Cycle: One example

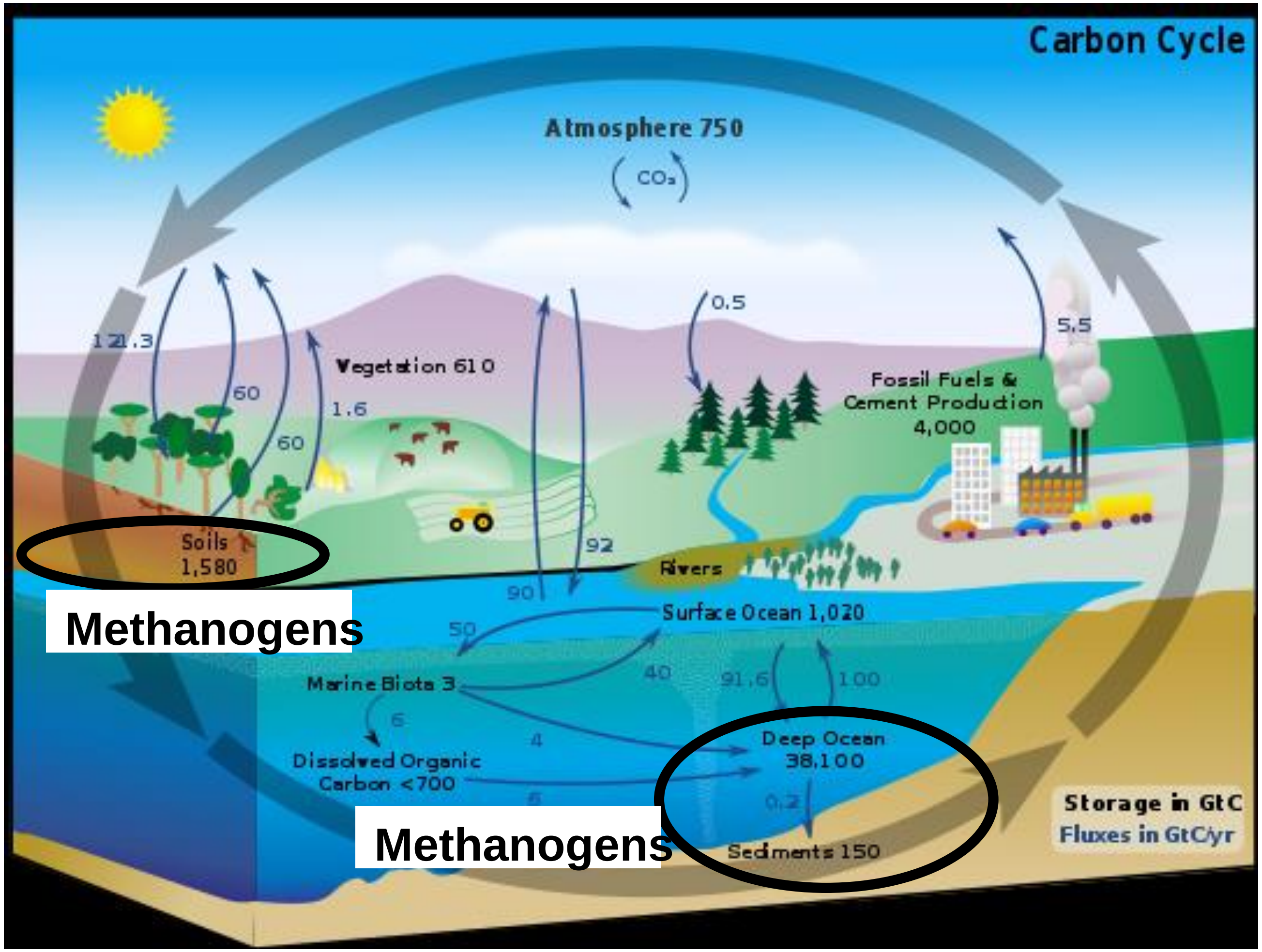
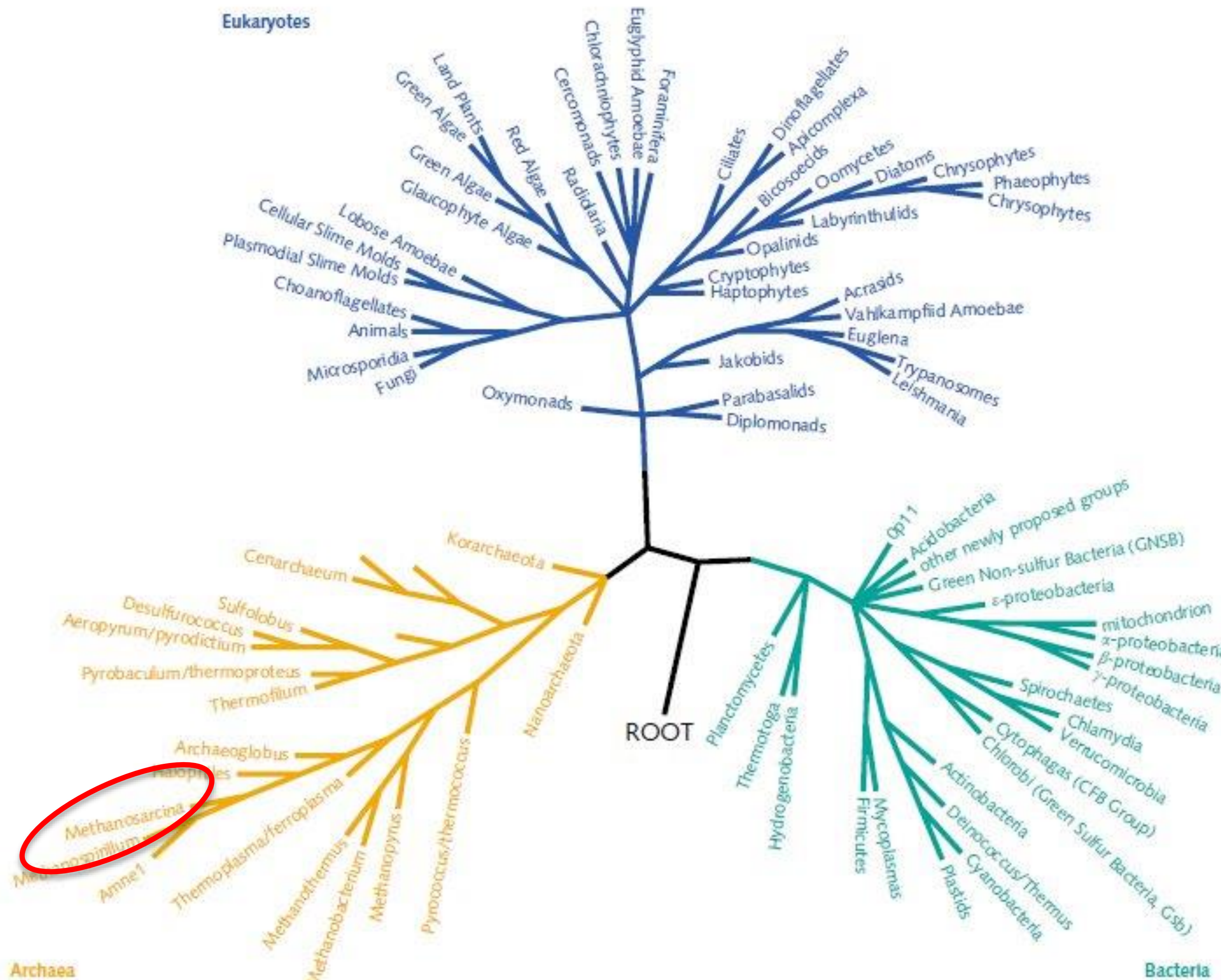


Image Source: NASA

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Metabolic network expansion around key species in phylogenic tree



stitute for
ns Biology
Revolutionizing Science. Enhancing Life.

Two keystone manual reconstructions

- *Methanosarcina barkeri*
  - Gonnerman…Metcalf, Price, *BMC Systems Biology,* 2012
  - Feist…Palsson, Ideker, *Molecular Systems Biology,* 2006

- *Methanosarcina acetovorans*
  - Benedict…Metcalf, Price, *Journal of Bacteriology*, 2011

The *M. acetivorans* metabolic network is highly curated

Number of:	
Metabolic Genes	746
Annotation Corrections	122 (16%)
Reactions*	757
Reactions with Genes	629 (83%)
Reactions with Literature	289 (38%)
Supporting Citations	159

*: Excluding exchanges

Benedict, Gonnerman, Metcalf, Price, *J. Bacteriol.* (2011)

Validation of the *M. acetivorans* metabolic model

Experimental

Simulated

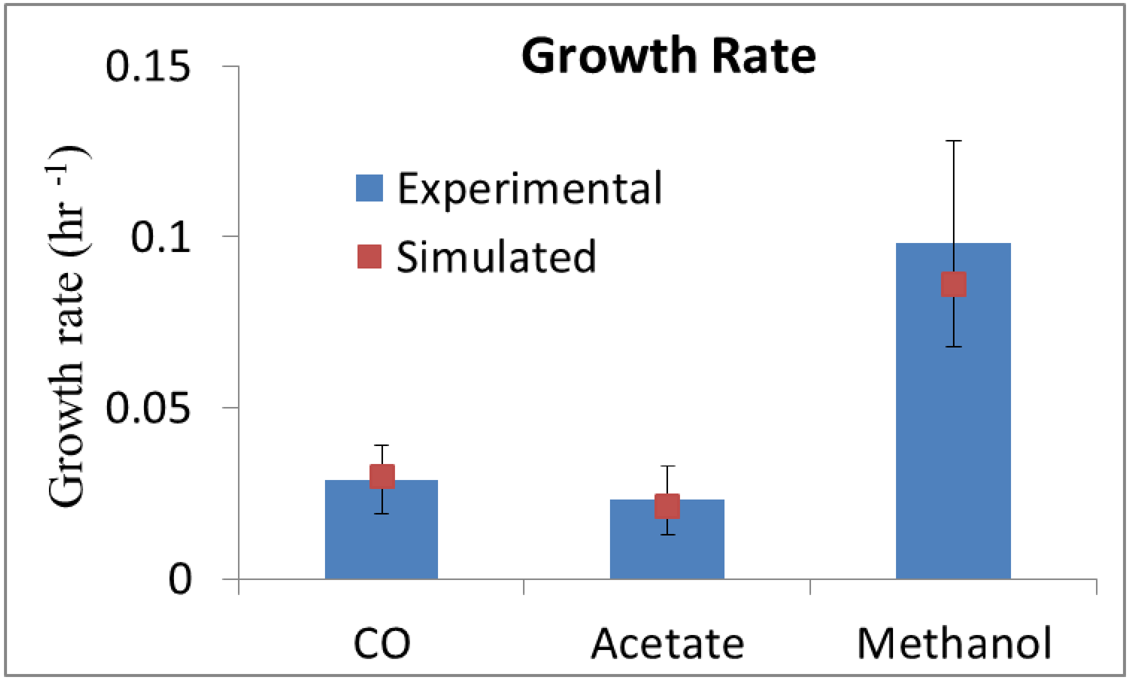
	Growth	No Growth
Growth	45/45	3/18
No growth	0/45	15/18
TOTAL	95% correct	

Gene knockout lethality

Benedict, Gonnerman, Metcalf, Price, *J. Bacteriol.* (2011)

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Quantitative Predictions of Growth Rates

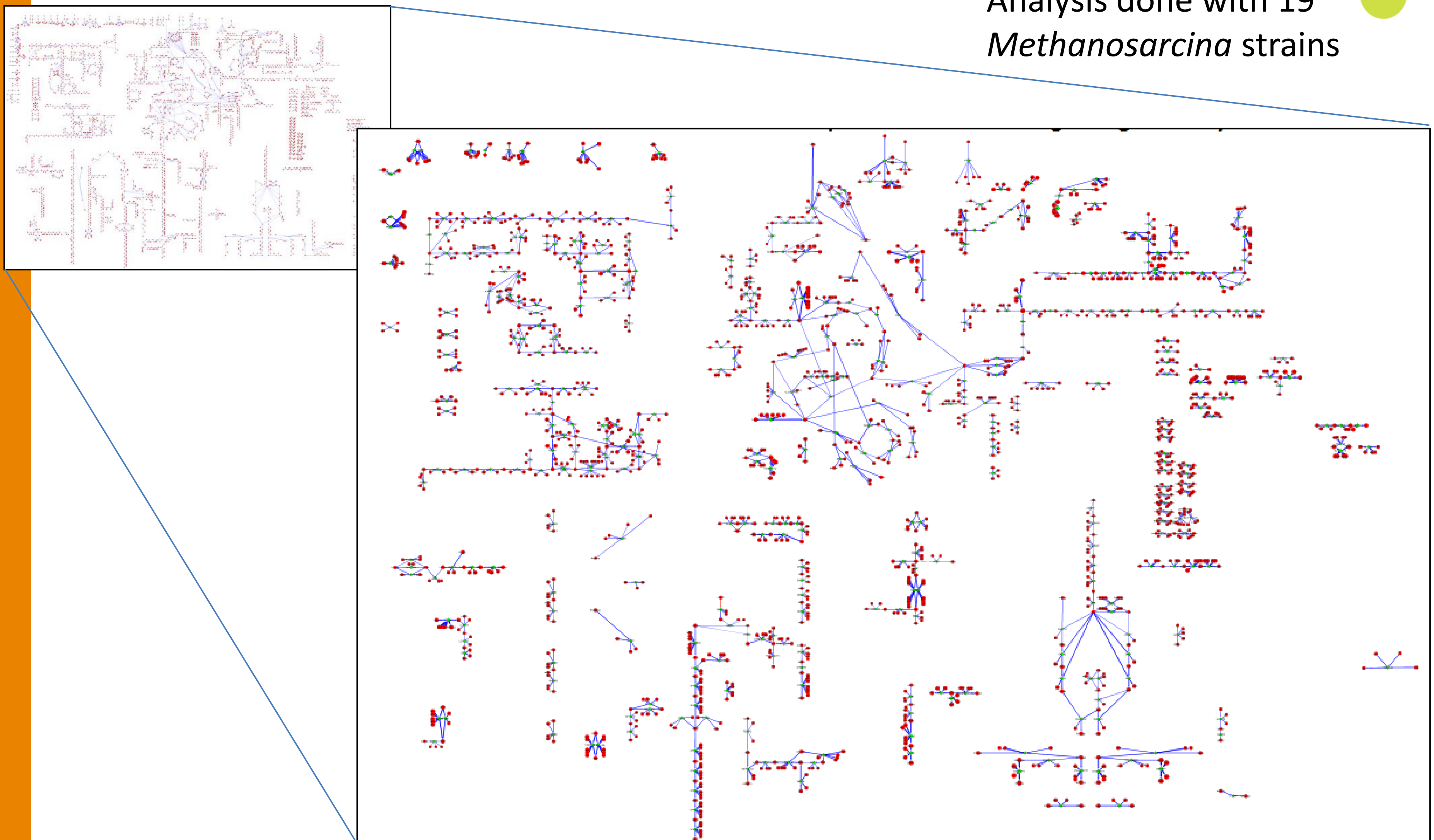


Growth rate predictions within experimental error

Benedict, Gonnerman, Metcalf, Price, *J. Bacteriol.* (2011)

Conserved *Methanosarcina* metabolism

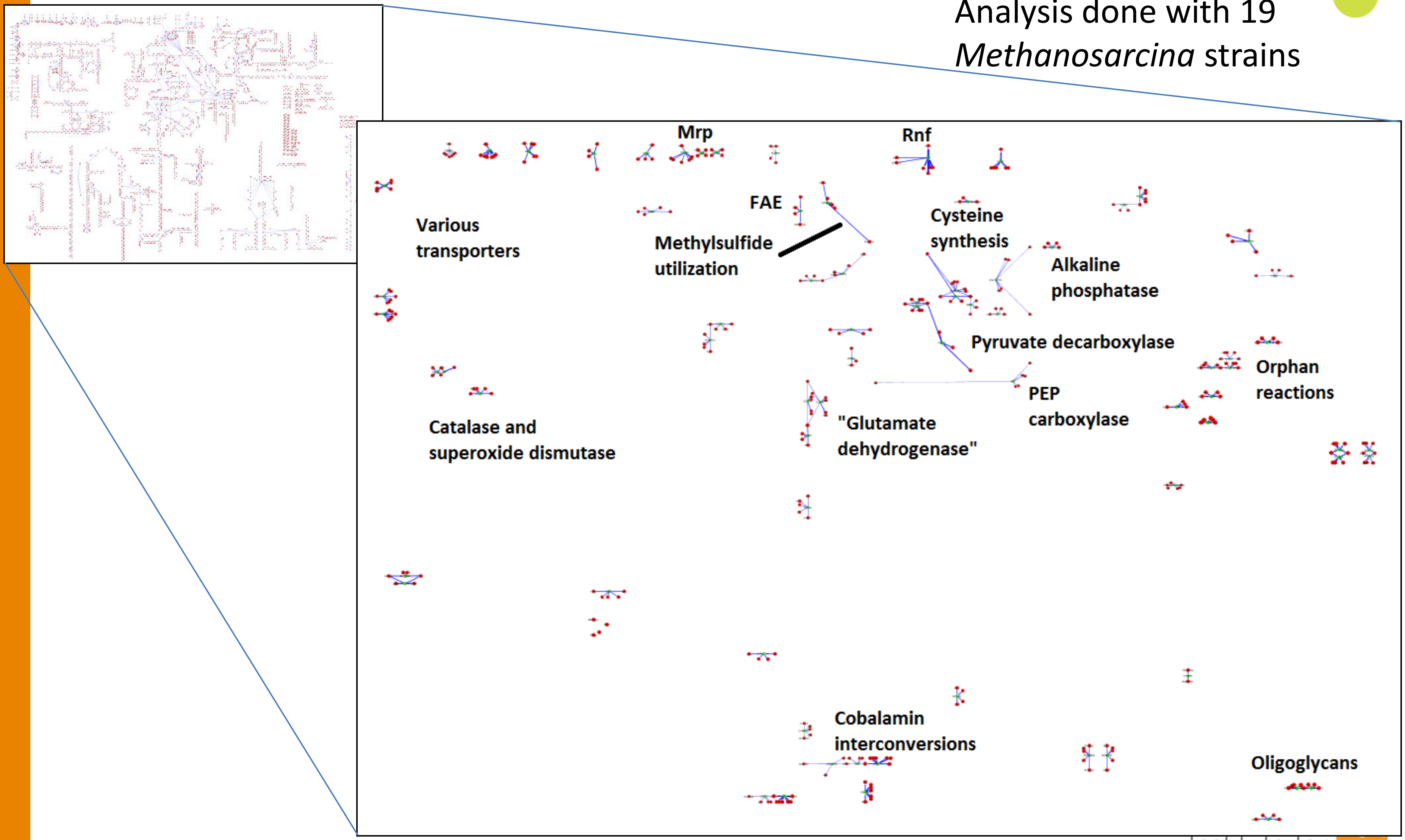
Analysis done with 19 *Methanosarcina* strains



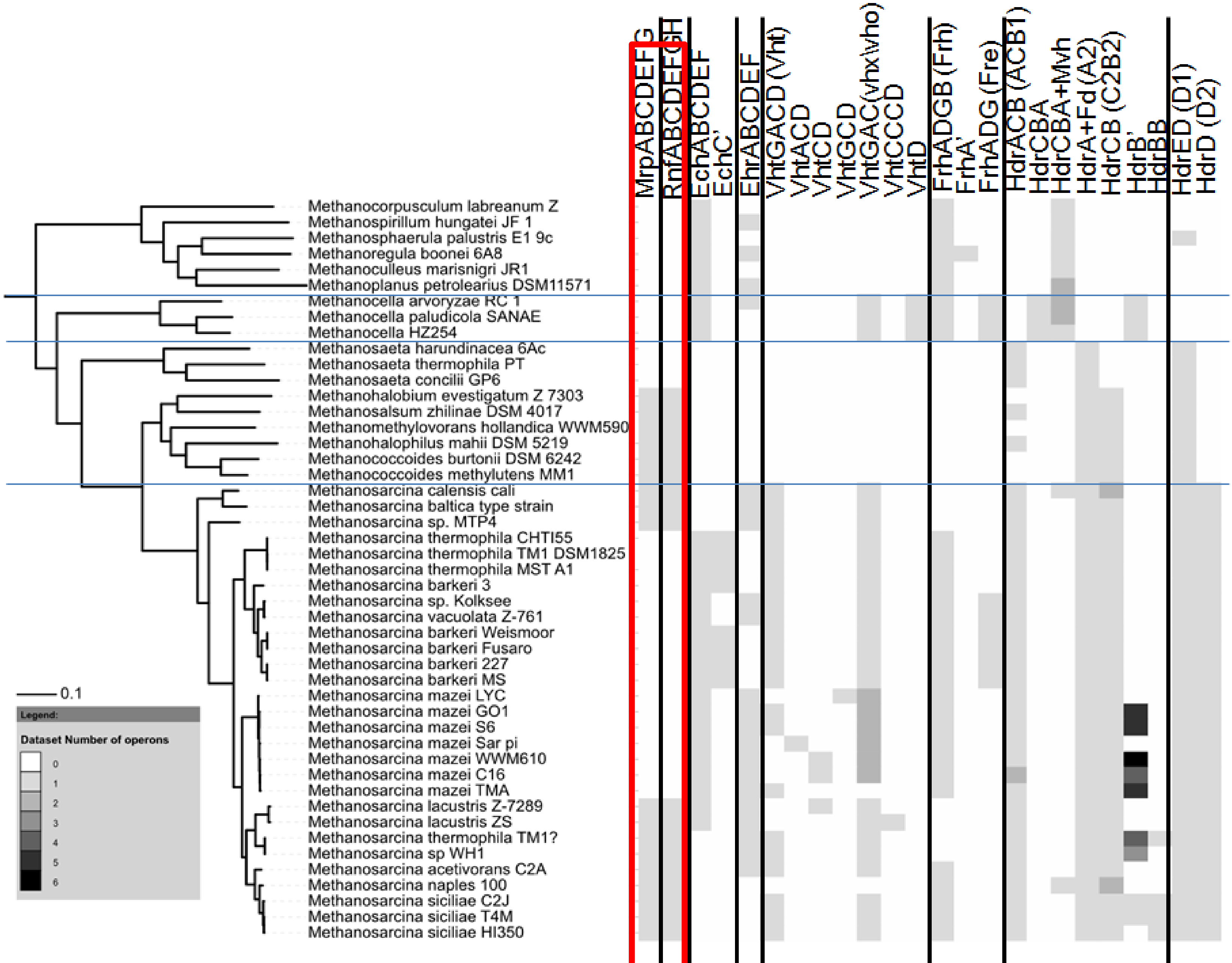
Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Variable *Methanosarcina* metabolism

Analysis done with 19 *Methanosarcina* strains

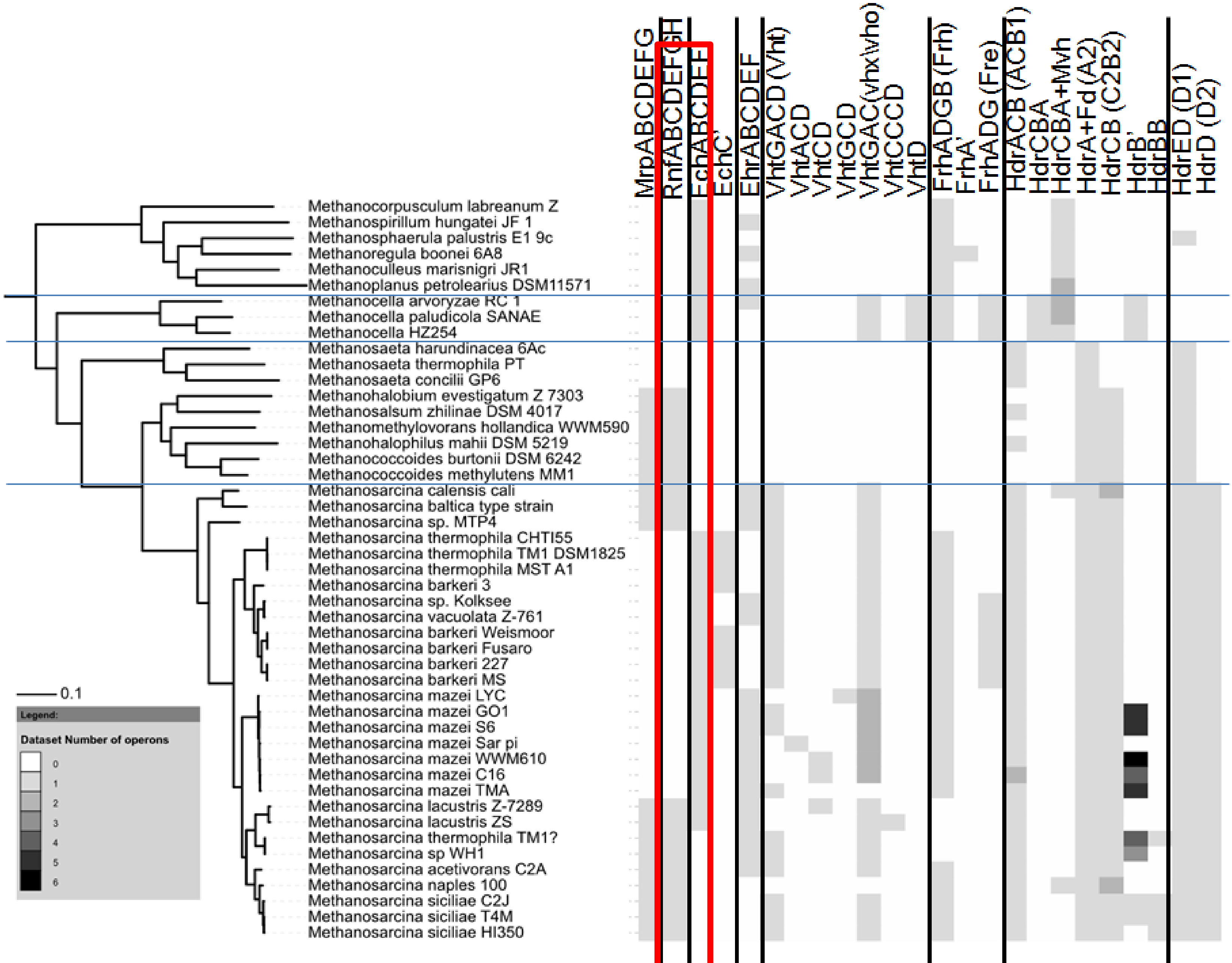


Institute for Systems Biology
Revolutionizing Science. Enhancing Life.



- Tree visualized using iTOL (http://itol.embl.de/).

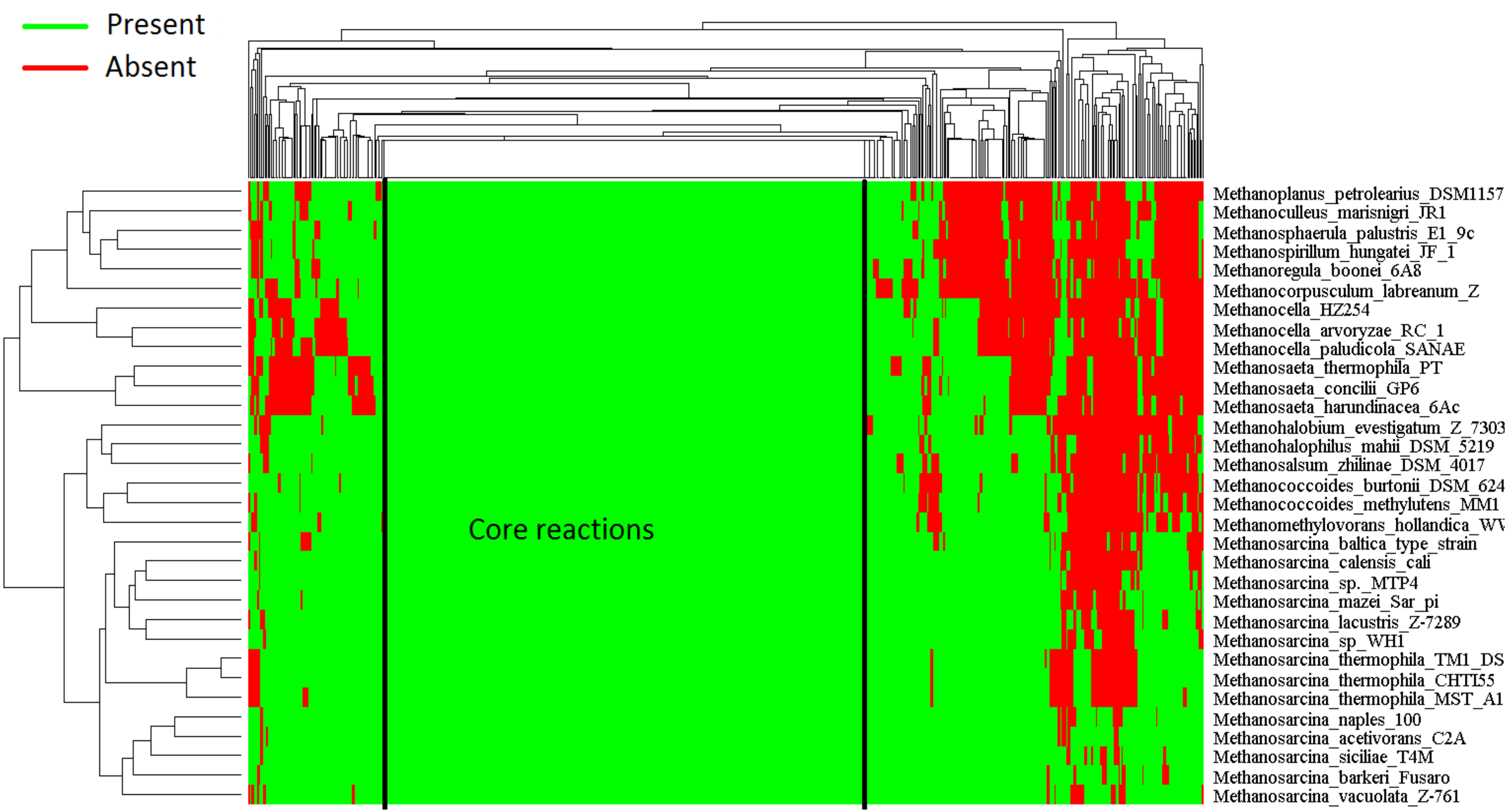
Institute for Systems Biology
Revolutionizing Science. Enhancing Life.



- Tree visualized using iTOL (http://itol.embl.de/).

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Comparison of curated *Methanosarcina*



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

ITEP: A toolkit for comparative genomics and curation

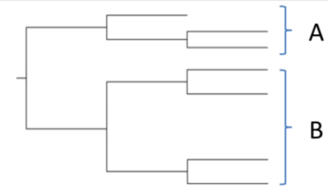
STEP 1:
Provide Source data

Annotated genomes

KBase PREDICTIVE BIOLOGY
DOE Systems Biology Knowledgebase

NCBI

Organism groupings



STEP 2:
Run startup scripts to build database

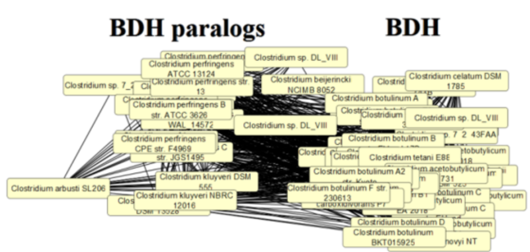
SQLite

Stores

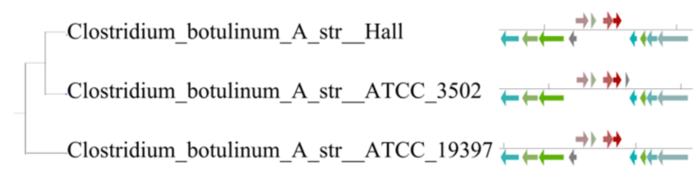
- Gene locations, sequences
- BLASTP, BLASTN, RPSBlast results
- Clustering results
- Contig sequences

STEP 3:
Use ITEP interfaces to the database to analyze and interpret genomes

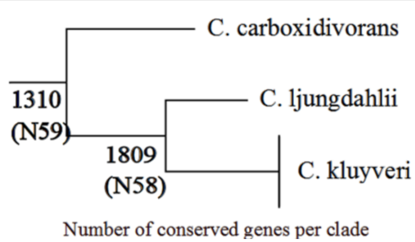
Cluster curation and visualization



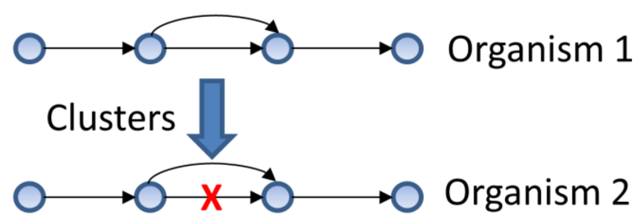
Phylogenies and gene context



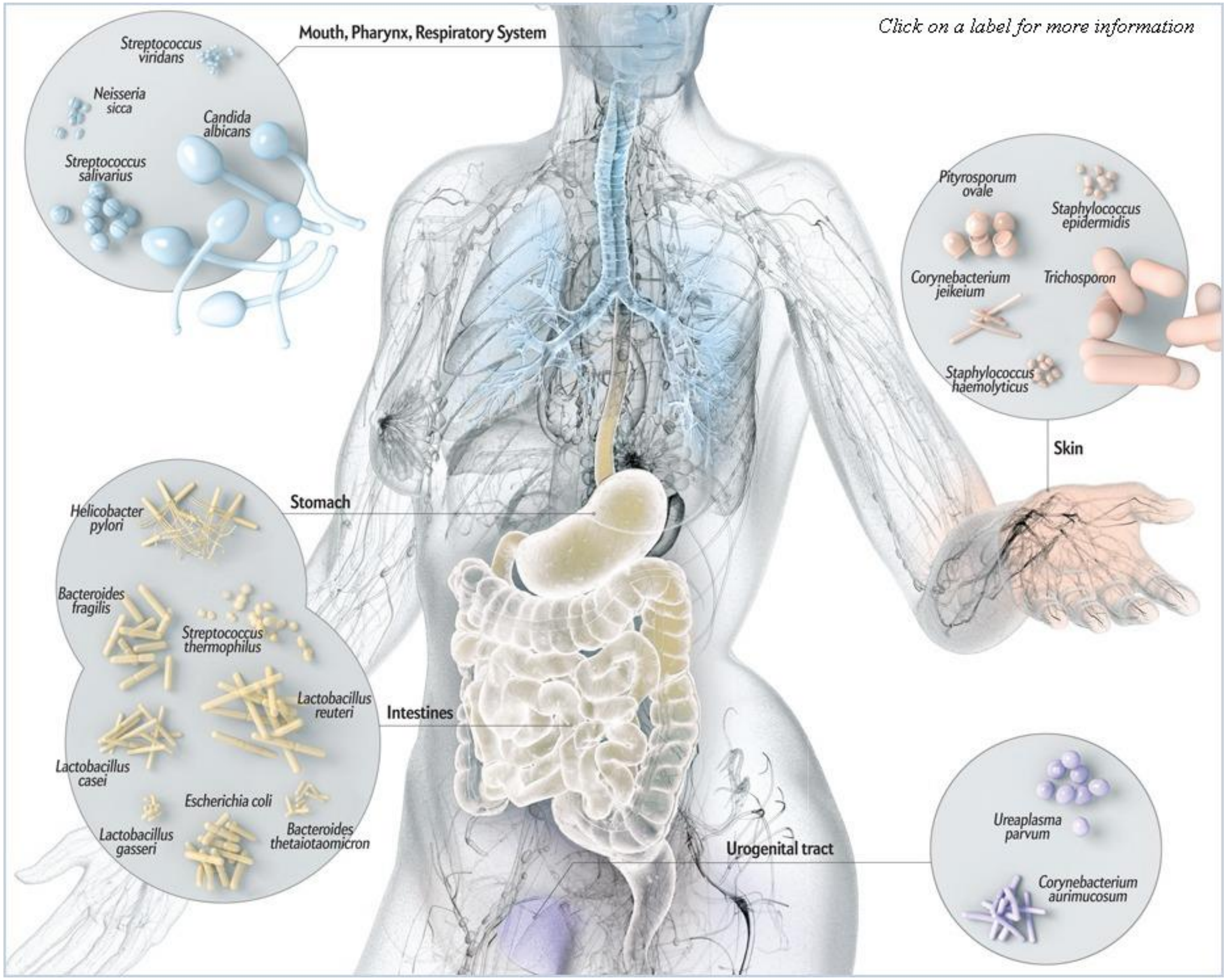
Analysis of gene gain and loss patterns



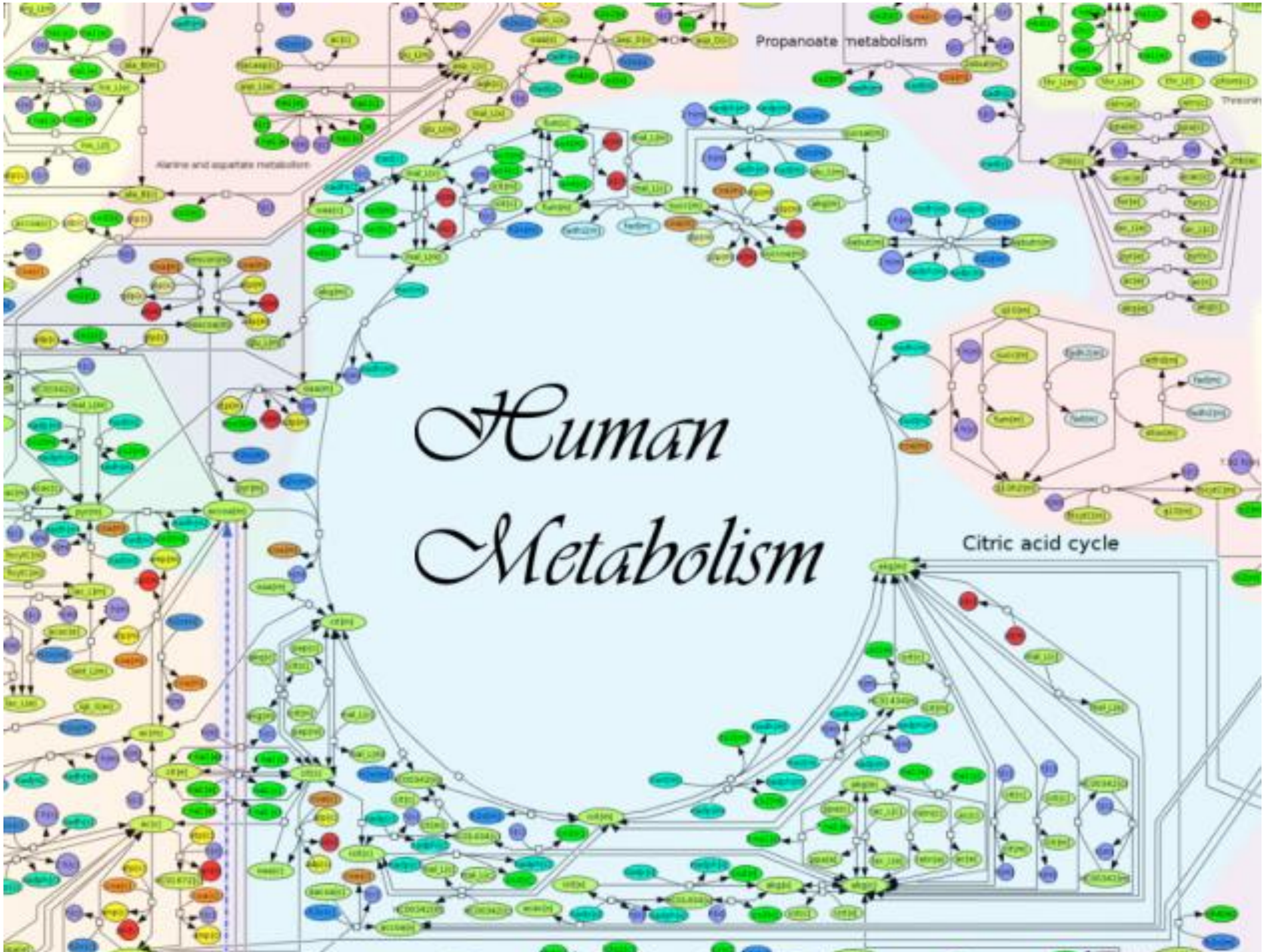
Metabolic reconstructions



Enormous scope of microbiomes motivates scalable approaches



A “Google Map” of Human Metabolism



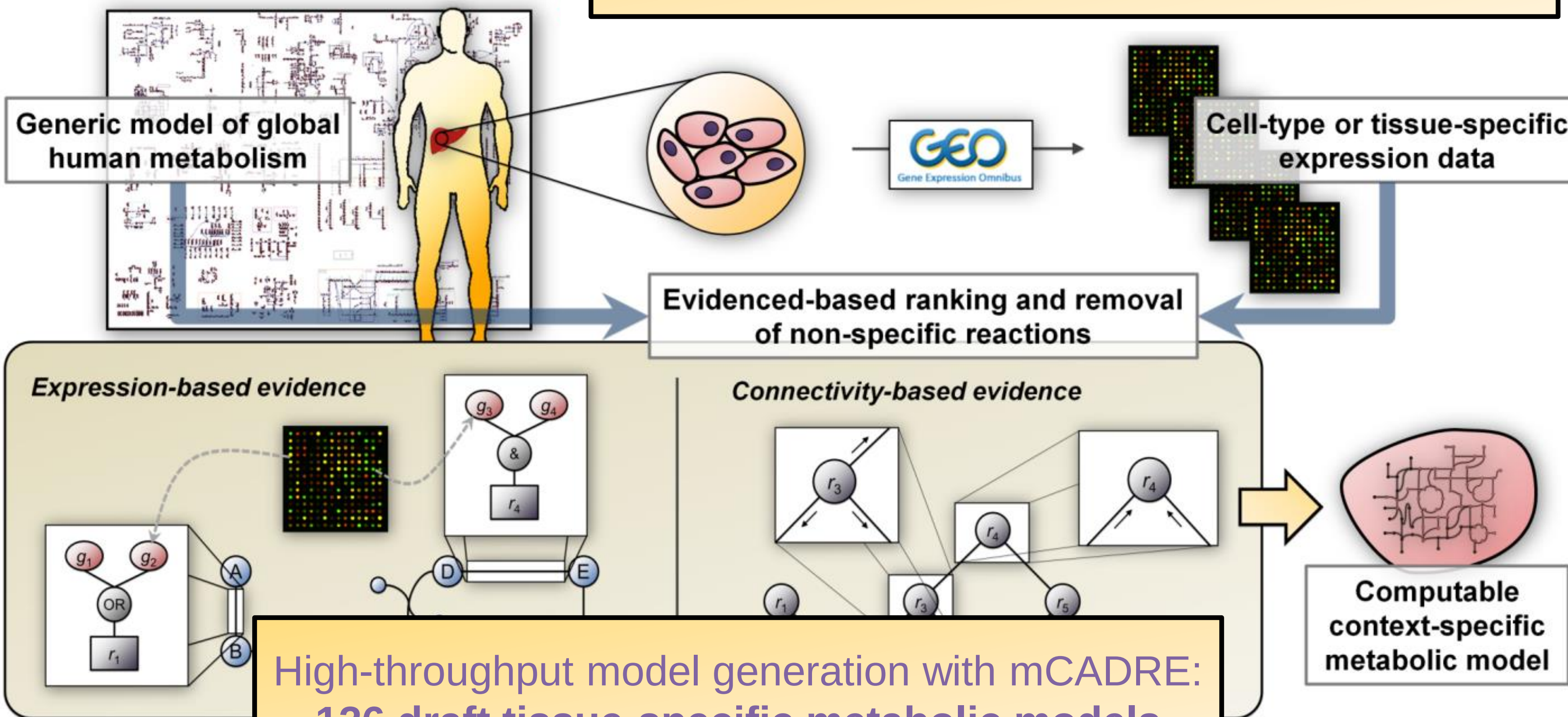
PROM: Chandrasekaran and Price, *PNAS*, 2010
mCADRE: Wang, Eddy, Price, *BMC Systems Biology*, 2012
GEMINI: Chandrasekaran and Price, *PLOS Computational Biology*, 2013
RECON2: Thiele et al, *Nature Biotechnology*, 2013

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

mCADRE

Semi-automated tissue-specific model generation

Metabolic Context-specificity Assessed by Deterministic Reaction Evaluation (mCADRE)



High-throughput model generation with mCADRE:
126 draft tissue-specific metabolic models

Wang, Eddy et al. *BMC Systems Biology* (2012)

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Early success: prediction and validation of synthetic lethal targets for combination therapy

LETTER

doi:10.1038/nature10363

Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase

Christian Frezza[1], Liang Zheng[1], Ori Folger[2], Kartik N. Rajagopalan[3], Elaine D. MacKenzie[1], Livnat Jerby[2], Massimo Micaroni[4], Barbara Chaneton[1], Julie Adam[5], Ann Hedley[1], Gabriela Kalna[1], Ian P. M. Tomlinson[6], Patrick J. Pollard[5], Dave G. Watson[7], Ralph J. Deberardinis[3], Tomer Shlomi[8]*, Eytan Ruppin[2,9]* & Eyal Gottlieb[1]

- Germline mutations of FH are responsible for hereditary leiomyomatosis and renal-cell cancer
- Use genome-scale metabolic model to explain how these cells survive without a complete TCA cycle
- Then identify specific target that is lethal in the new metabolism, and NOT to normal human metabolism
- Cell experiments then validated the finding

Nature 2011

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Take home points 1

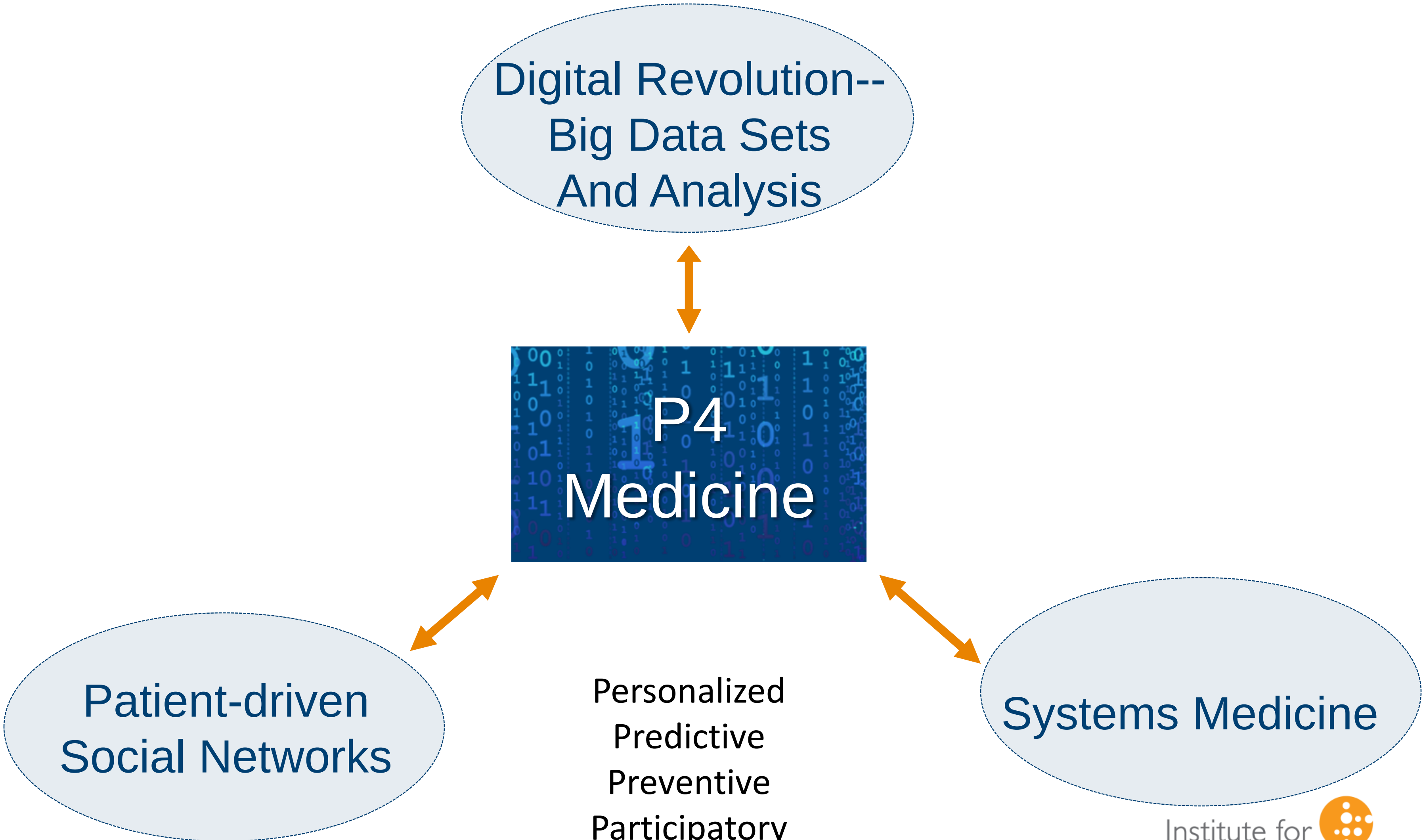
- One critical factor in understanding the environment is mapping out the functions of microbes
- Genome-scale metabolic models and community interaction network models can aid in this task
- Similar approaches can be employed in humans, providing a platform for studying symbiotic and pathogenic relationships
- Initial success for human metabolic network modeling in predicting new candidate cancer therapy that has been initially experimentally validated

Bringing P4 Medicine to Practice

Institute for Systems Biology

Revolutionizing Science. Enhancing Life.

The convergence of three revolutions leads to proactive P4 medicine

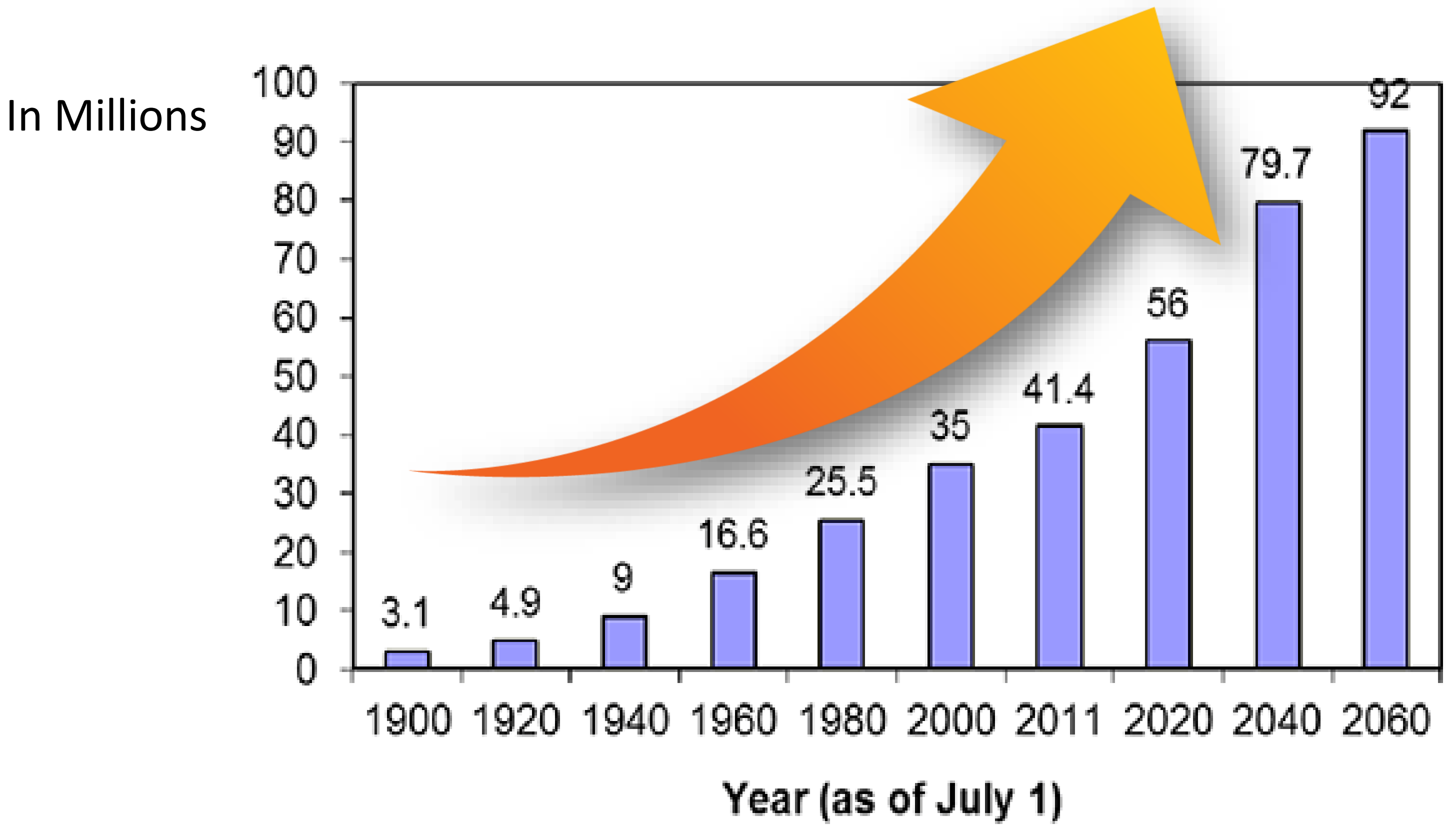


Institute for Systems Biology
Revolutionizing Science. Enhancing Life.


CARDIAC
Science

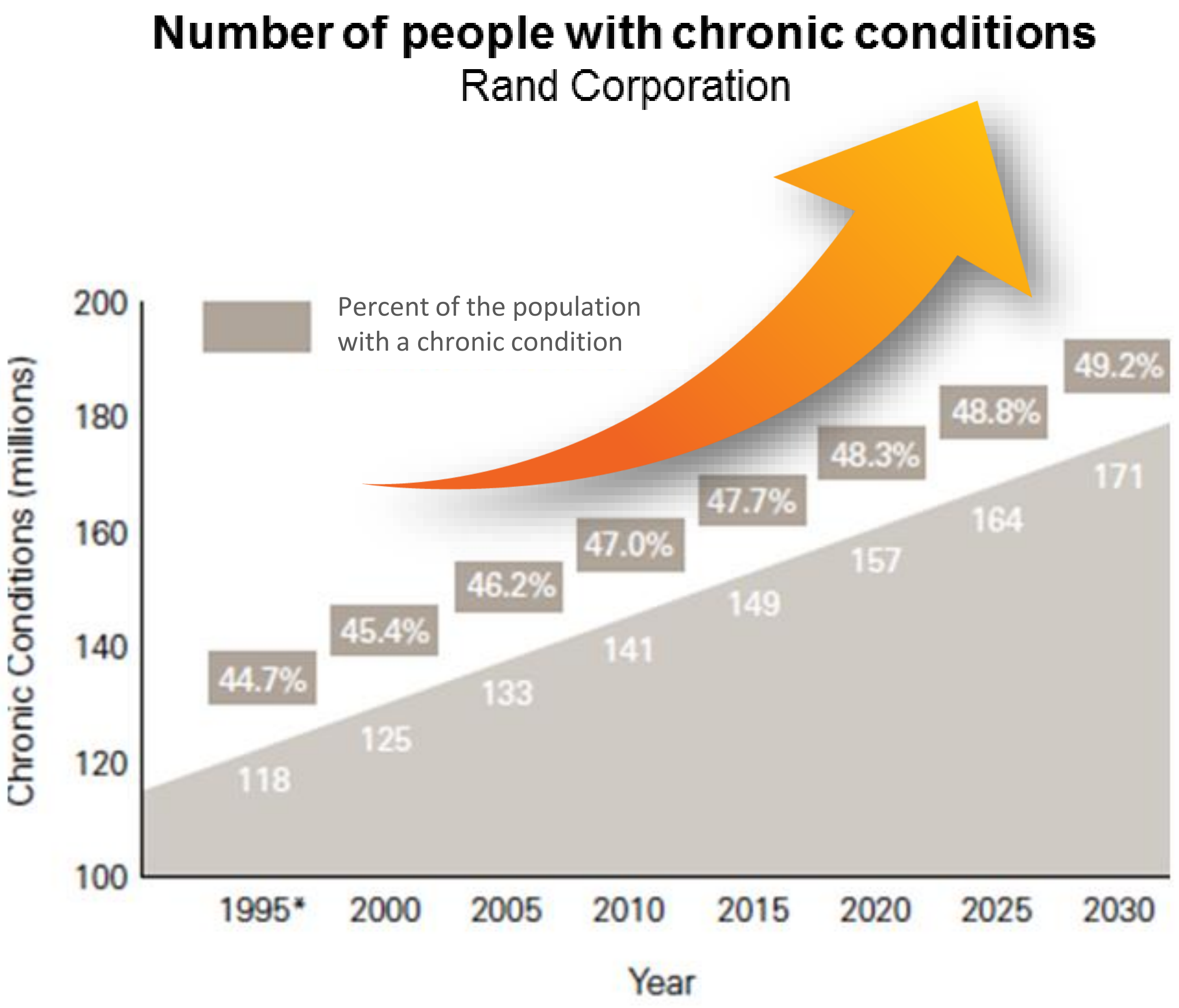
Aging Populations: Skyrocketing

Number of persons 65+ (1900-2060)
Administration of Aging



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Chronic Disease: Skyrocketing



PIONEER 100
Hundred Person Wellness Project

The Problem

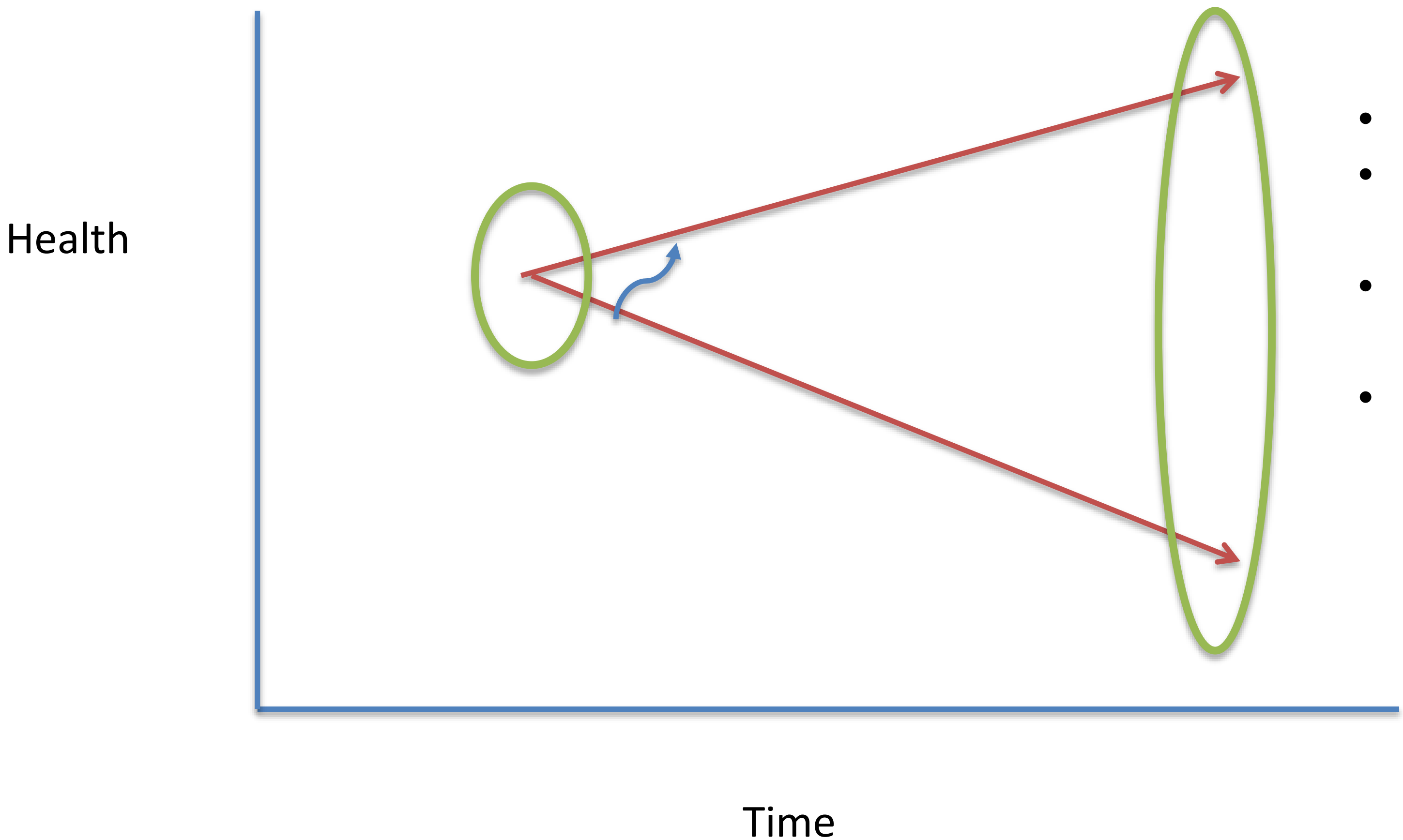
- Current medical field focused entirely on care after illness
- Data is based on p…
- …know when we may become chronically ill

What if this were different?

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

PIONEER 100
Hundred Person Wellness Project

Health: What do we really want to understand?



- Bigger effect size
- Larger likelihood for paper
- Smaller sample size required
- Less $

- **Longitudinal**
- More complex study
- Larger sample size required
- More $$$

PIONEER 100

Hundred Person Wellness Project

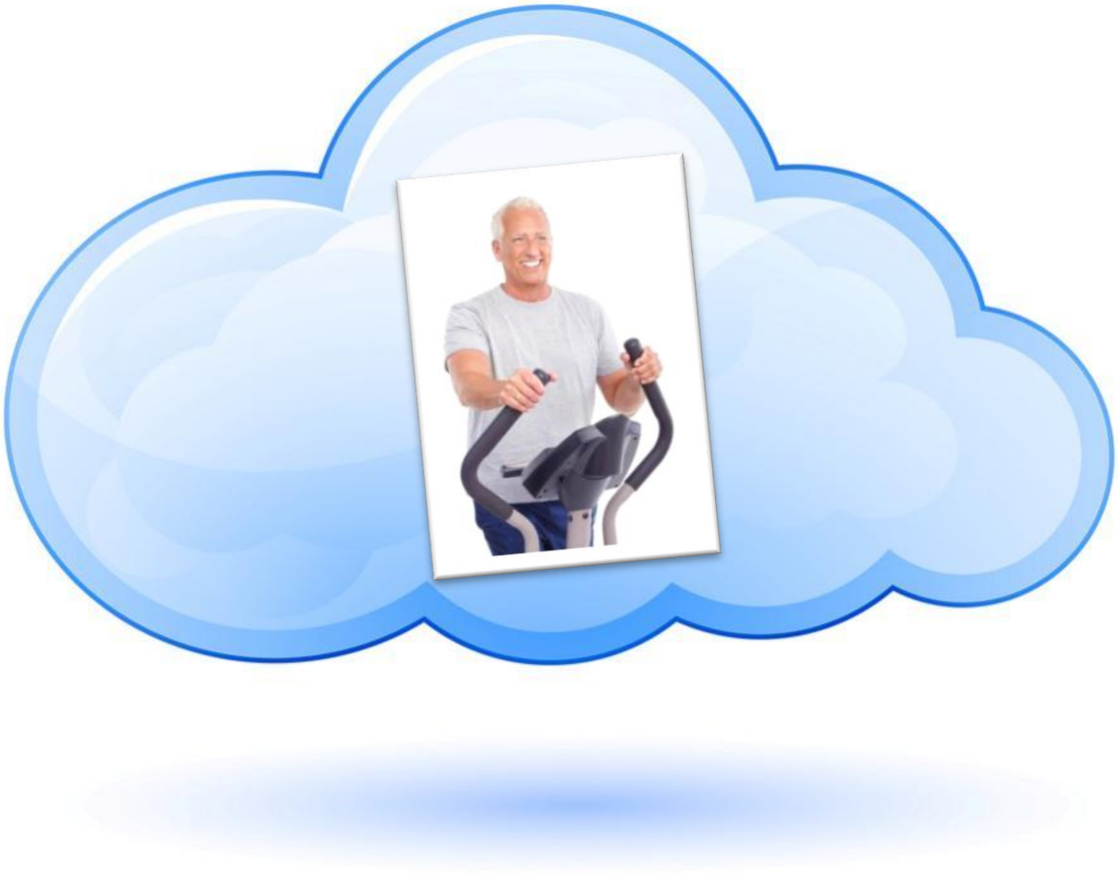
- 100 participants
- 9-month study started in March
- IRB approved

- Whole genome sequence
- Detailed blood, urine, saliva measurements 3x
- Gut microbiome 3x
- Continual activity and lifestyle monitoring
- Discovery research on samples
- Data integration
- Coaching, events, education

A Systems Medicine Approach to Wellness

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

A Unique, N=1 Data Cloud

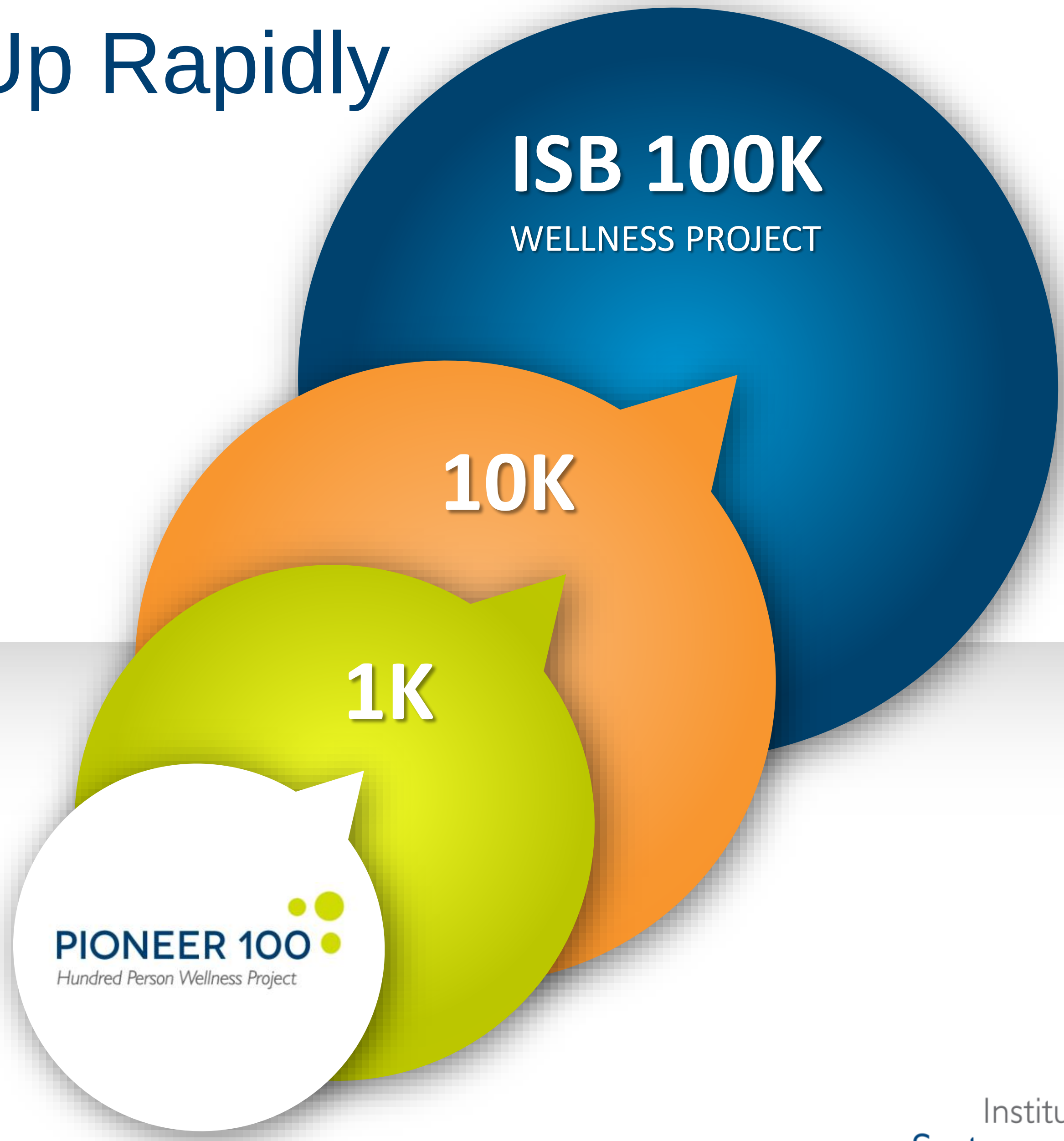


Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

… Times 100

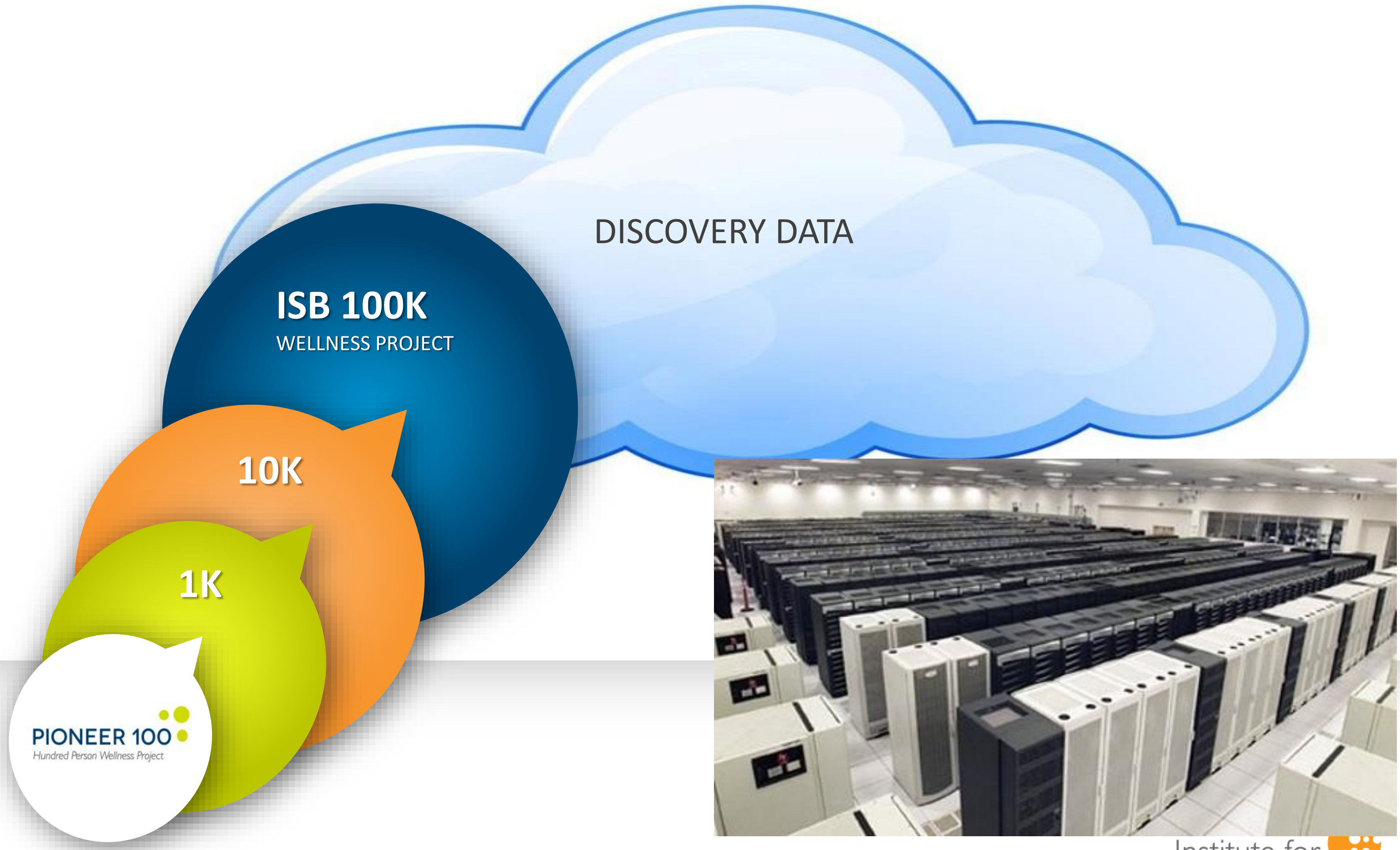
Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Scaling Up Rapidly



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Sharing for discovery



Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Three scales of analysis

- **Short term**—optimize wellness and reduce disease for each individual patient and reduce the costs

N = 1

- **Intermediate term**—longitudinal assessments of individuals over time and what brings about change

N = 1 to tens (families/friends)

- **Long term**—generate a data base from individuals that will allow us to follow transitions from wellness to disease for major diseases

N <= 100K

From complexity, simplicity

> “I would not give a fig for the simplicity on this side of complexity; but I would give my right arm for the simplicity on the far side of complexity”
>
> Oliver Wendell Holmes

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Goals

1. Establish scientifically validated **metrics** for wellness
2. Determine **"actionable items"** to present to individuals participating in 100-person pilot project
3. Identify **transitions** between disease and wellness
4. Identify what **benefits** are compelling for individuals

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Conclusion and Further Reading

Clinical OMICs INNOVATOR

Promoting Wellness & Demystifying Disease:

The 100K Project

Leroy Hood, M.D., Ph.D., and Nathan D. Price, Ph.D.

Hood and Price, *Clinical Omics,* May. 2014

EDITORIAL

SYSTEMS BIOLOGY

Demystifying Disease, Democratizing Health Care

Leroy Hood is President of the Institute for Systems Biology, Seattle, WA 98109, USA. E-mail: lhood@systemsbiology.org

Nathan D. Price is Associate Director of the Institute for Systems Biology, Seattle, WA 98109, USA. E-mail: nprice@systemsbiology.org

Citation:
L. Hood, N. D. Price, Demystifying disease, democratizing health care. *Sci. Transl. Med.* **6**, 225ed5 (2014).

UNSUSTAINABLE COST INCREASES THREATEN THE GLOBAL HEALTH CARE SYSTEM, and further progress is stymied more by societal than technological factors. Only by engaging health care consumers (that is, patients) as pioneers who provide both health-related data and insights into pathophysiology can we meet these societal challenges and thus accelerate the pace of biomedical innovation.

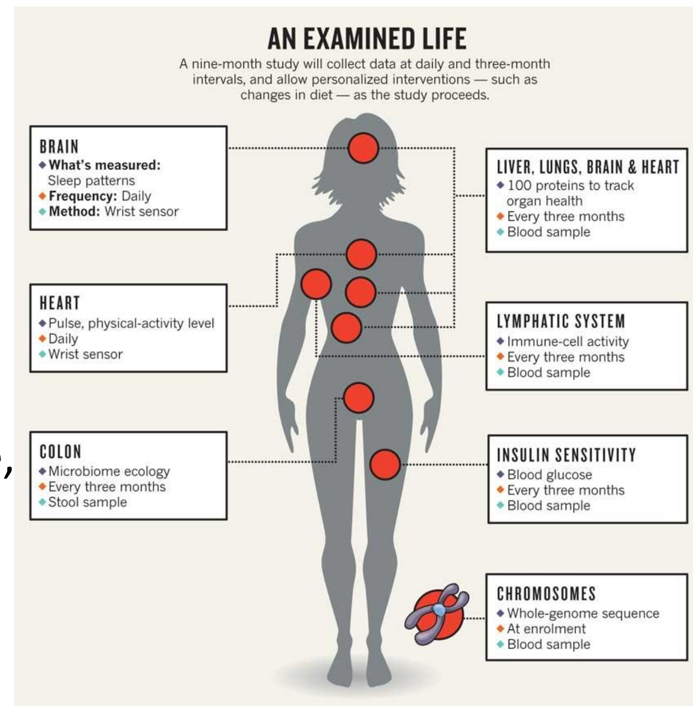
In March 2014, the Institute for Systems Biology will launch a longitudinal, Framingham-like study (www.framinghamheartstudy.org) of 100,000 (100K) healthy individuals that we believe will be instrumental in bringing predictive, preventive, personalized, and participatory (P4) medicine to patients. Participatory medicine means that patients, researchers, physicians, and the entire health care community join forces to transform the practice of medicine to make it more proactive than reactive—and, in turn, less expensive and more effective (*1*).

PEOPLE POWER

A systems approach is necessary for the effective management of complex diseases (*1*). This fundamental component of P4 medicine is built on two central features. The first is a conviction that, in 5 to 10 years, each patient will have a dynamic data cloud consisting of billions of diverse types of data points and that medicine will be informed by computational analyses that reduce high-dimensional data to actionable hypotheses designed with the intent of optimizing wellness and minimizing disease in individual patients. The second feature is that integration of patient data will reveal biological networks that specify health and are altered in disease, and that through an understanding of these differences, one can gain fundamental insights into disease mechanisms. Such insights are essential for developing more effective diagnostic and therapeutic approaches. Indeed, such an approach has already provided powerful new technologies and strategies (*2*) that have brought us to the brink of P4 medicine (*3*).

At its foundation, P4 medicine is about quantifying wellness and demystifying disease. Individual data clouds will let us predict future wellness and disease. The preventive element focuses on how well we can improve individual wellness and take actions to stop or de-

Nature, News piece, Feb. 2014



Hood and Price, *Science Translational Medicine* (2014)

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.

Nathan D. Price Research Laboratory Institute for Systems Biology

Scientific Project Manager
Julie Bletz

Postdoctoral Fellows
Seth Ament
Nicholas Chia *(Now Asst. Prof. at Mayo Clinic)*
James Eddy
Cory Funk
Ben Heavner
Saheed Imam
Kyung Kwa Kim
Younhee Ko
(now at Yonsei *University*)
Vineet Sangar
Vangelis Simeonidis

Graduate Students
Matthew Benedict
Sriram Chandrasekaran
(now Harvard Junior Fellow)
John Earls
Piyush Labhsetwar
Shuyi Ma
Andrew Magis
Matthew Richards
Chunjing Wang
Yuliang Wang
(now Res. *Asst. Prof. at OHSU*)

Visiting Researchers
Sascha Schäuble, PhD (Friedrich Schiller University-Jena)
Zhuo "Joy" Wang, PhD (Shanghai Jiao Tong University)
Carl Chang, PhD (Chang Gung Memorial Hospital)
Juan Zhang, PhD (Jiangnan University)

Collaborators
Lee Hood (ISB) **Bill Metcalf** (Illinois)

Funding
- National Institutes of Health
- Department of Energy
- National Science Foundation
- CHDI
- Bill & Melinda Gates Foundation
- Grand Duchy of Luxembourg
- Department of Defense
- Energy Biosciences Institute (BP)
- Roy J. Carver Young Investigator Award
- Camille Dreyfus Teacher-Scholar Award

Energy Biosciences Institute
Roy J. Carver CHARITABLE TRUST
The Camille & Henry Dreyfus Foundation

Institute for Systems Biology
Revolutionizing Science. Enhancing Life.