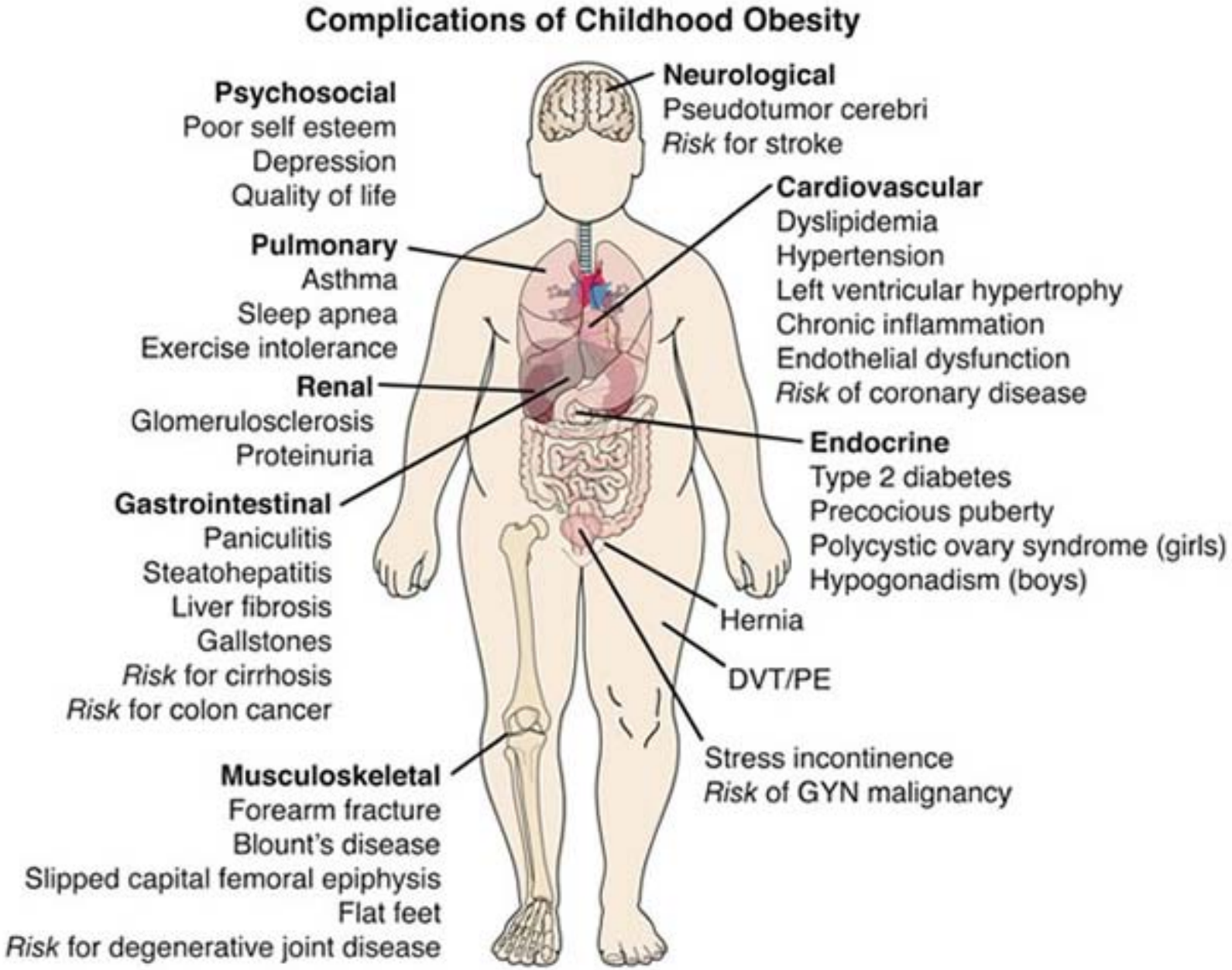
Mechanism-associated Phenotypes for Genetic Analyses (MAPGen)

Proposed Roadmap Initiative for 2011

Gail Weinmann, MD
Division of Lung Diseases, NHLBI
November 21, 2008
Council of Councils



Source: www.obesityhelp.com

PROBLEM

Focus on organ-based manifestations results in:

- Lack of focus on shared underlying mechanisms (difficulty leveraging research across institutes)
- Inadequate/inaccurate phenotypes for genotype associations
- Missed opportunities in therapeutic pipeline

Mechanism-Associated Phenotypes for Genetic Analysis

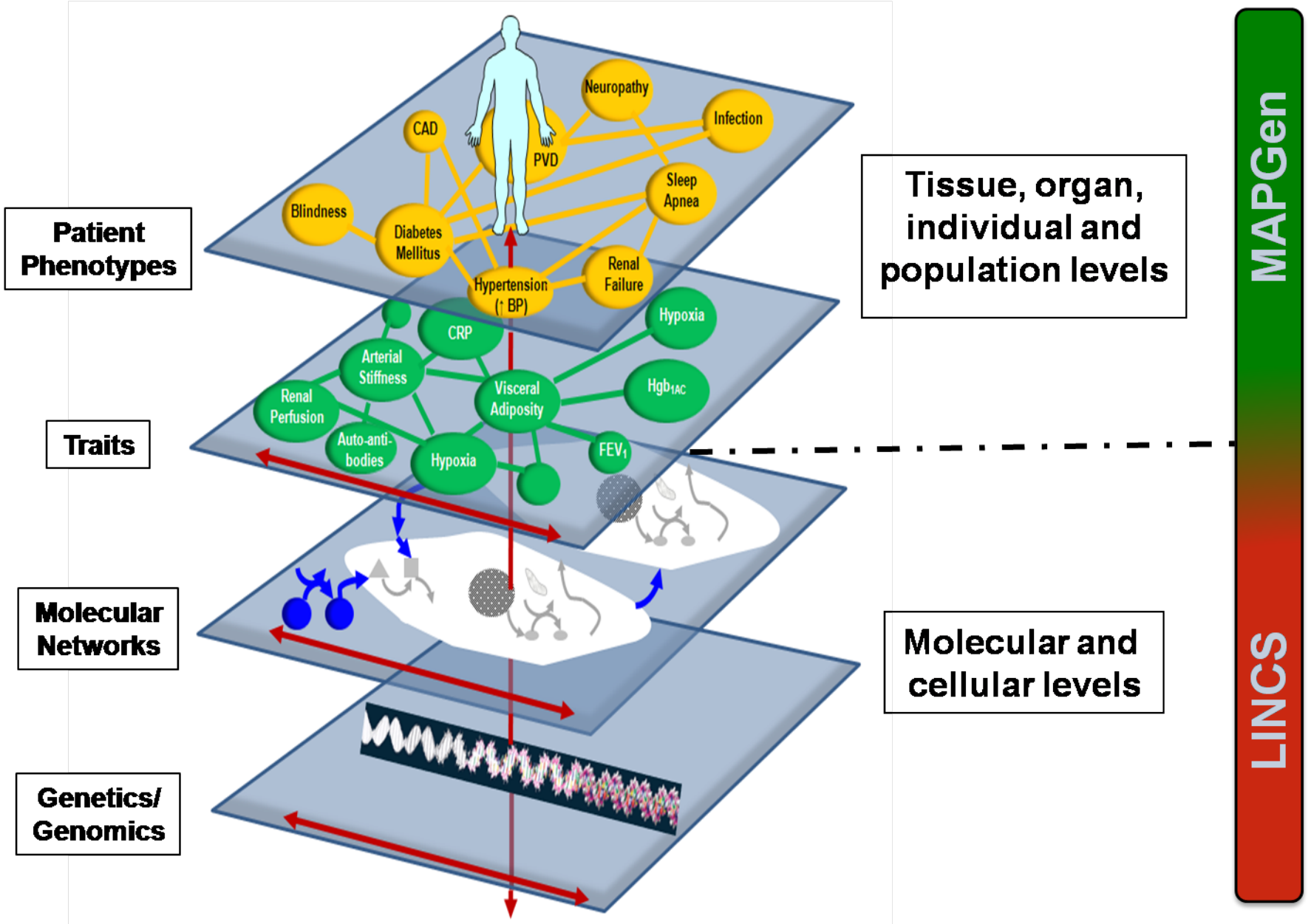


Figure adapted from Barabasi, *New Engl J Med* 357:404-7 (2007)

MAPGen Concept

Transformative concept -Define human illnesses based upon the underlying pathobiology.

Goal – Redefine disease phenotypes according to shared mechanisms.

Product – evolving knowledge of phenotypes and associated tools based upon shared mechanisms for use in research.

Diseases are interconnected.

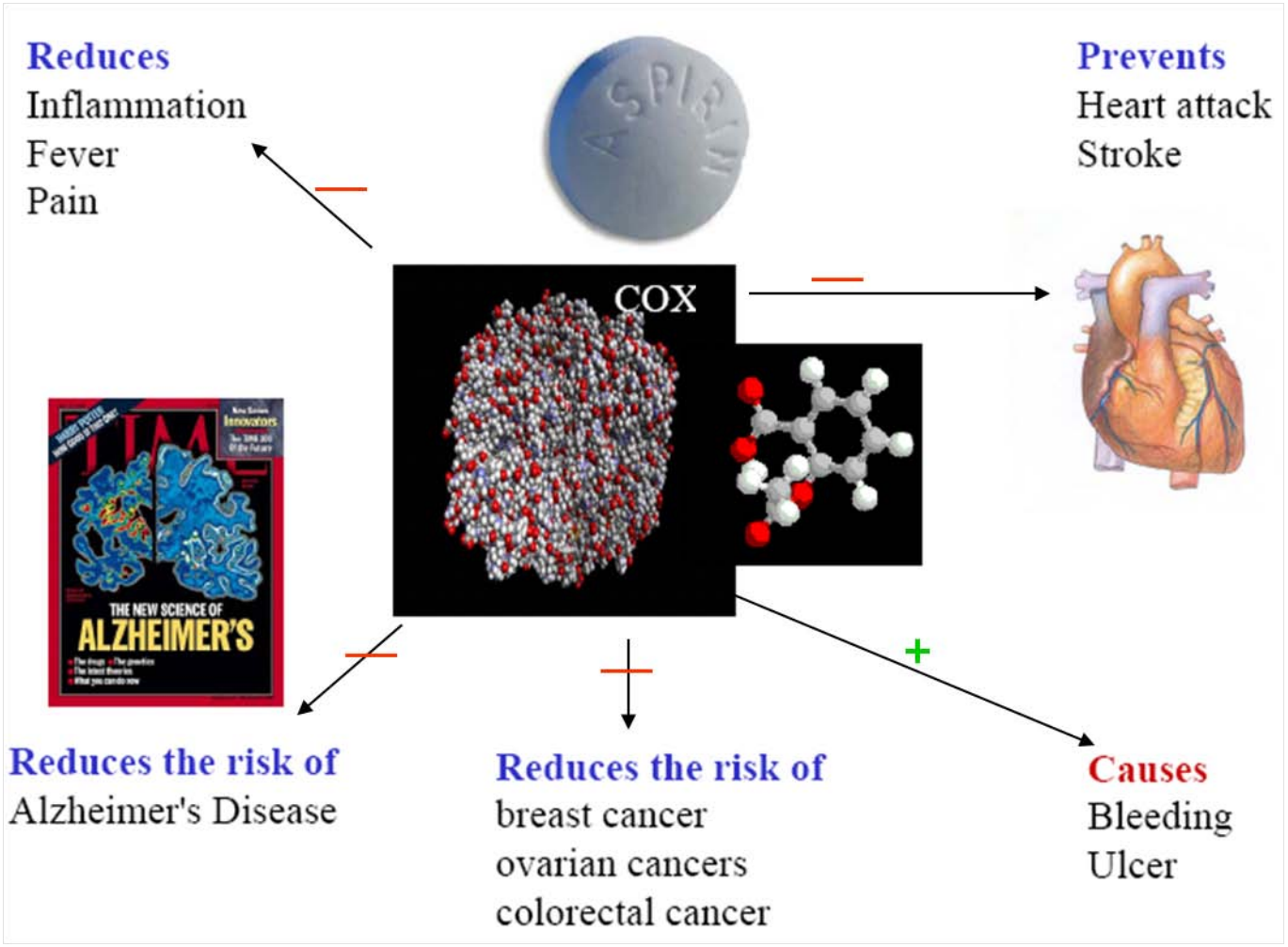
Supporting observations include:

- ✓ Many genes to one disease
- ✓ One gene to many diseases
- ✓ Many to many: interconnections among biological pathways.

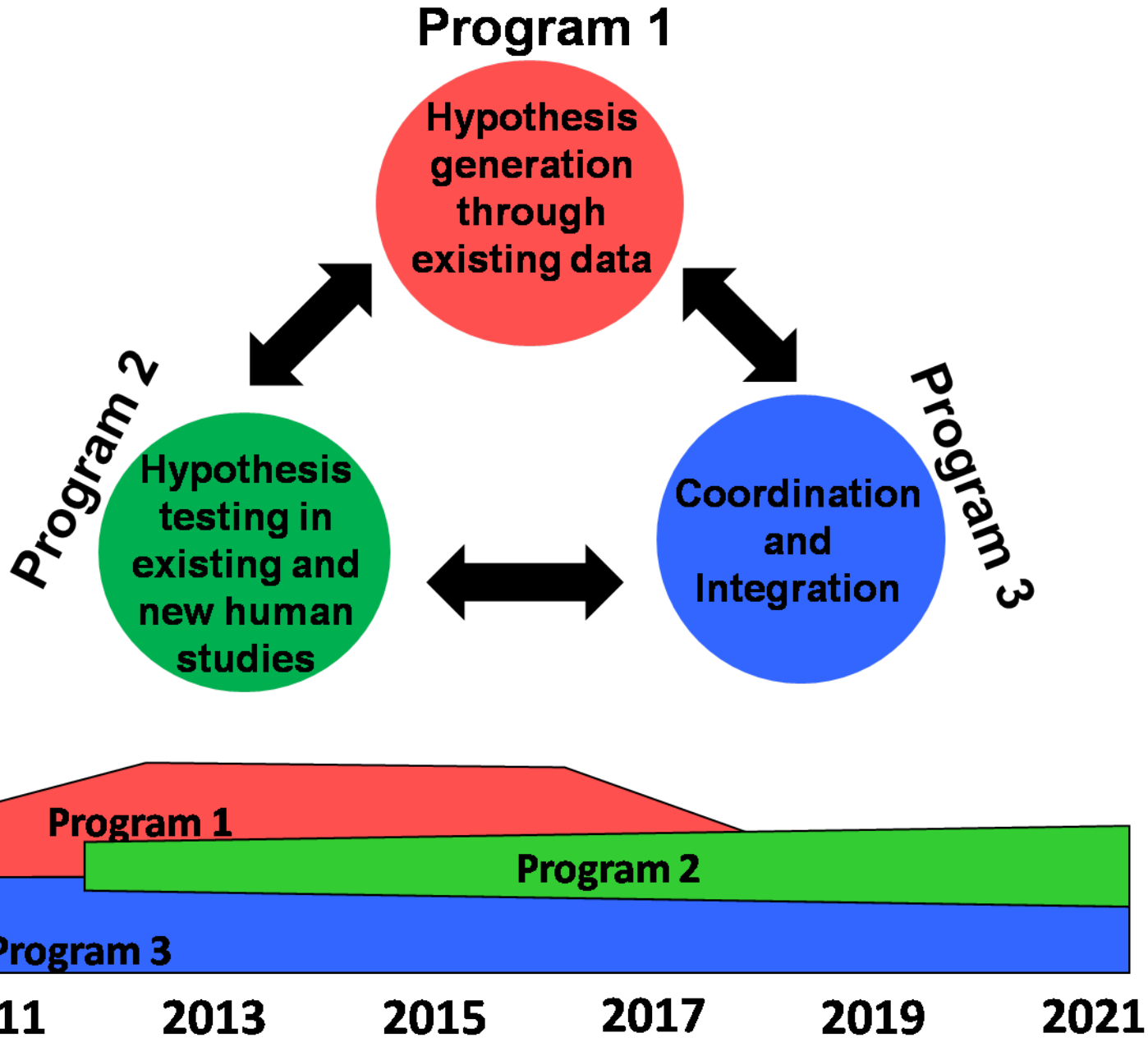
A variant associated with nicotine dependence, lung cancer and peripheral arterial disease

Nature **452**, 638-642 (3 April 2008)

One drug to many organs



Overview of MAPGen



Program 1: Existing large data sets permit insights

Search 1.5 million patient records on 161 disorders stored for pragmatic purposes (such as billing) for co-morbidity, correcting for age, sex, age-at-onset, and ethnicity

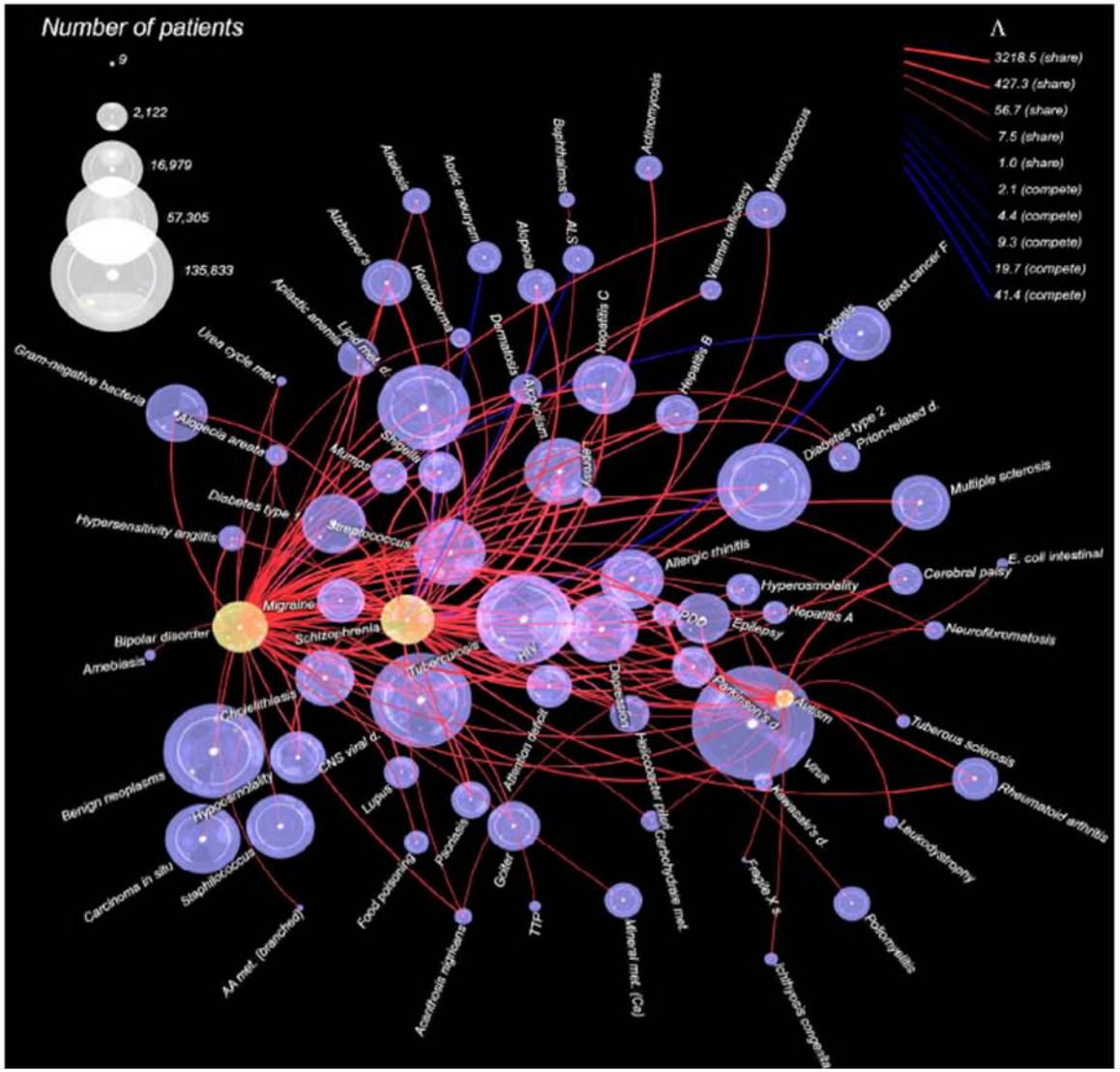


Fig. 3. Significant correlations (that we interpret as genetic overlap) among three neurodevelopmental disorders (autism, bipolar disorder, and schizophrenia; corresponding nodes are shown in yellow) and all other disorders in our data set (blue nodes). The volume of each sphere (disease) is proportional to the number of patient records annotated with the corresponding phenotype, as explained in the key. The arcs represent significant correlations among phenotypes, with negative correlations shown in blue and positive correlations shown in red. Thicker arcs represent stronger correlations; see key.

Andrey Rzhetsky, *et al.* (2007) Proc. Natl. Acad. Sci. U. S. A. 104, 11694–11699

Program 1 – Hypothesis generation through screening for interconnections using existing data and samples

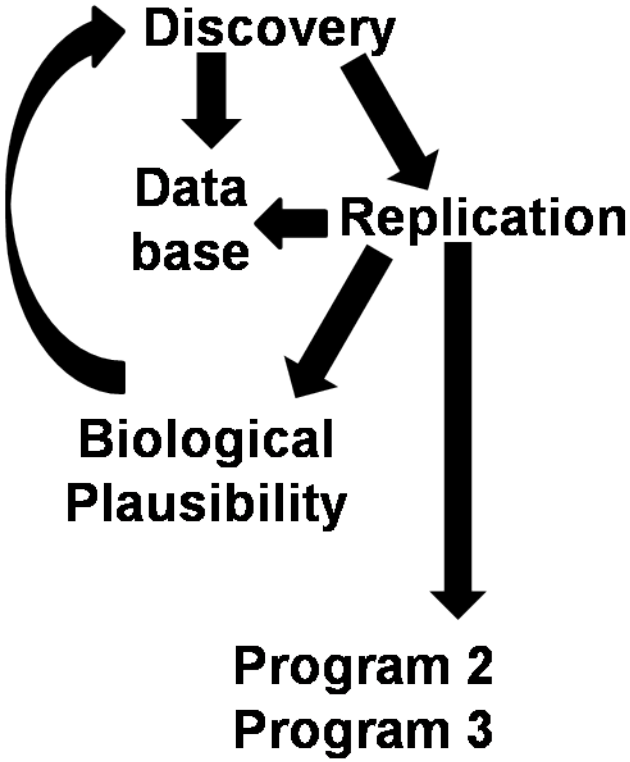
Resources

- Medical literature
- Electronic medical records
- Characterized cohorts (clinical, molecular, and GWAS)
- Stored biosamples
- Phenotype databases
- Program 2
- Program 3

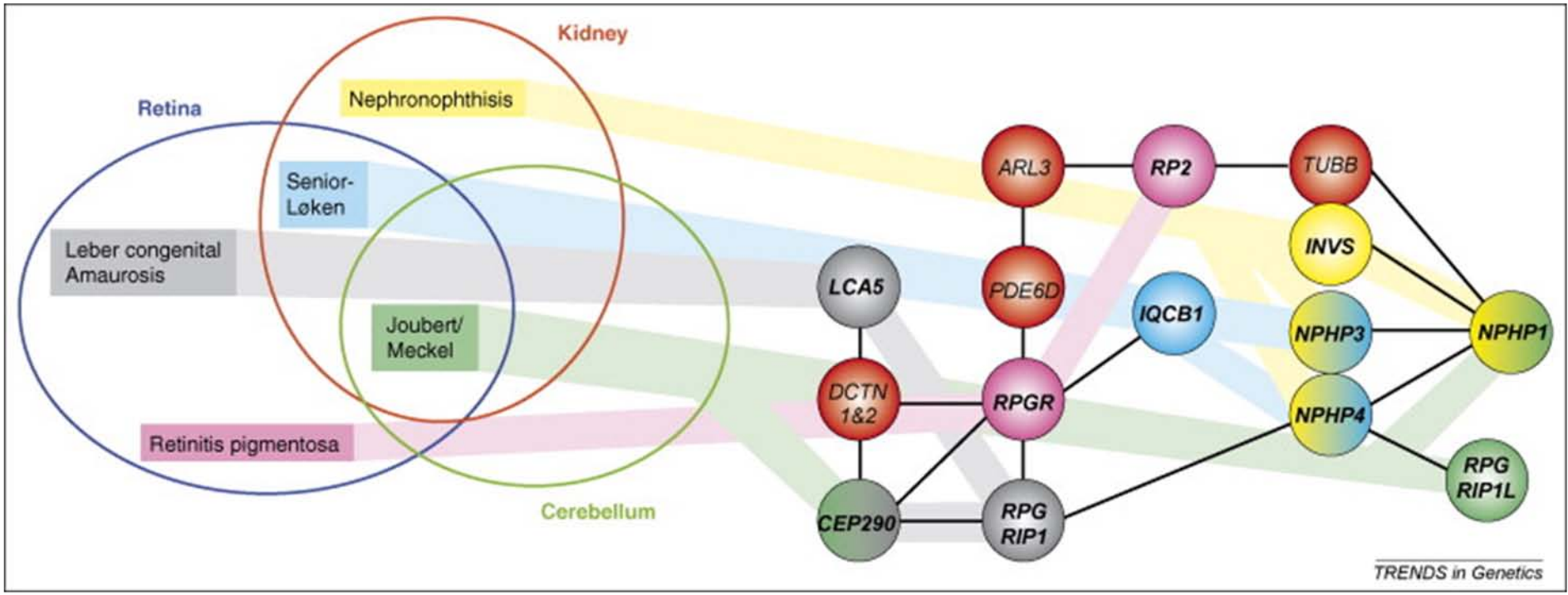
Approach

- Literature mining
- Data mining
- Meta-analysis
- Lab analyses
- Integration of clinical data, molecular phenotyping, genotyping

Process



Program 2: Understanding shared pathobiology reveals new insights.



Recognizing overlapping phenotypes (cystic kidneys, retinal degeneration, polydactyly, and brain malformations) resulting from mutations in ciliary proteins will permit identification of potential ciliopathies.

Martin Oti, *et al.* Trends in Genetics Vol. 24 No.3

Program 2 – Hypothesis testing through analysis of existing or new human studies

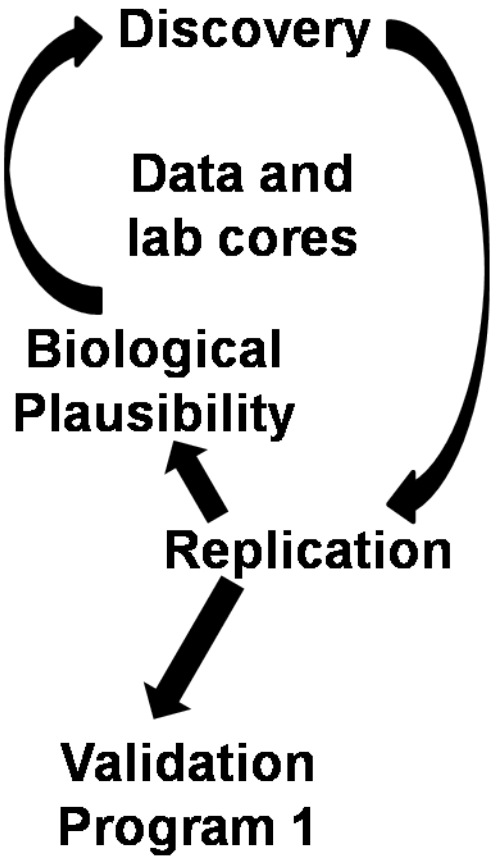
Resources

- Program 1
- Stored biosamples with associated clinical data
- Ongoing clinical studies
- Program 3
- LINCS

Approach

- Elucidation of markers and mechanisms across diseases
- Testing of hypotheses in: clinical, laboratory, and genetic findings across diseases

Process



Program 3 – Program Coordination and Data Integration

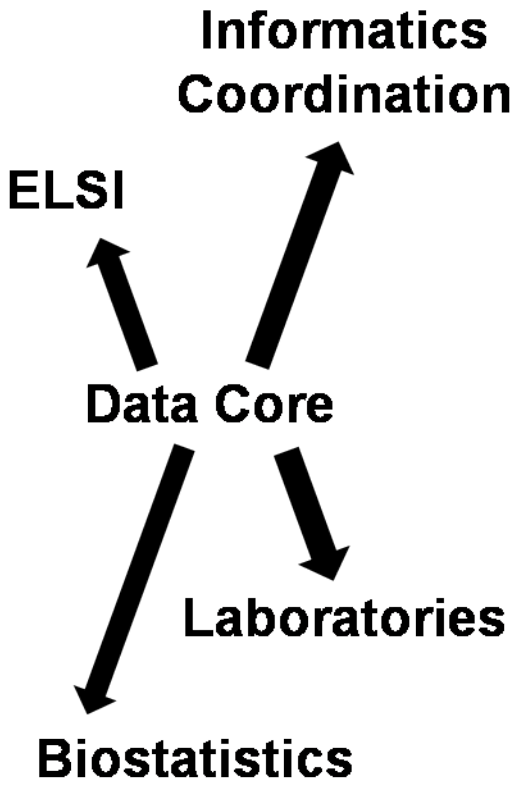
Resources

- Coordinating center
- Biostatistical expertise
- Bioinformatics & Computational science expertise
- Laboratory cores
- Ethical, legal, social implication expertise
- CTSAs

Approach

- Informatics coordination and exchange
- Development and dissemination of statistical tools
- Full capture and integration of databases
- Quality control
- Development and housing of dynamic web portal

Process



MAPGen Working Group Members

NHLBI

Croxton, Thomas

Gan, Weiniu

Kiley, James

Larkin, Jennie

Levy, Daniel

McDonald, Cheryl

Michelson, Alan

O'Donnell, Christopher

Punturieri, Antonello

Weinmann, Gail

NIBIB

James Luo

NIAAA

Guo, Q. Max

Hereld, Dale

NIAID

Fenton, Matthew

NIDA

Conway, Kevin

NICHD

Moye, John

NIDDK

Star, Robert

NIGMS

Remington, Karin

Lyster, Peter

NIMH

Lehner, Thomas

Little, A. Roger

NINDS

Riddle, Robert

Tagle, Danilo

NCRR

Olga Brazhnik

THANK YOU

Sample Research Programs

Program 1 - Hypothesis generation

- **Screen medical literature for associations**
- **Screen electronic medical records for associations**
- **Screen NIH supported databases**

Program 2 - Hypothesis Testing

- **Propose and validate common role of pathway across traditional diseases**
- **Propose and validate mechanisms of treatments potentially effective across diseases**
- **Differential organ response to environmental stressor**

Program 3 - Support

- **Integrate databases**
- **Develop statistical tools for use in screening large, uneven data sets**
- **Develop dynamic web portal**
- **Develop and evaluate statistical methods for evaluating phenotype-genotype association**

Sample Research Programs

Program 1 - Hypothesis generation

- Screen medical literature for associations
- Screen electronic medical records for associations
- Screen NIH supported databases

Program 2 - Hypothesis Testing

- Propose and validate common role of pathway across traditional diseases
- Propose and validate mechanisms of treatments potentially effective across diseases
- Differential organ response to environmental stressor

Program 3 - Support

- Integrate databases
- Develop statistical tools for use in screening large, uneven data sets
- Develop dynamic web portal
- Develop and evaluate statistical methods for evaluating phenotype-genotype association