Towards a yeast reference interactome

Alice Desbuleux1,2, Tiziana Cagarelli1, Yang Wang1, Thomas Rolland1, Noor Jaiikhani1, Yun Shen1, Katja Luck1, Kerstin Spirohn1, Sudharshan Rangarajan1, Charles Ogagan Jr.1, Jean-Claude Twizere2, Mike Calderwood1, Tong Hao1, Yves Jacob1,3, David Hill1, Quan Zhong1,4, Marc Vidal1, Benoit Charloateux1

1DFCI Center for Cancer Systems Biology (CCSB) and Department of Cancer Biology, and Harvard Medical School, Department of Genetics, Boston, MA; 2Pompeu Fabra University, Institute of Biomechanics of Catalonia, Barcelona, Spain; 3Inserm, U772, Unité de Biologie Cellulaire et de Génétique Cellulaire, Paris-Cachan, France; 4Wright State University, Dayton, OH

Despite an increasing number of interactomic datasets already available for model organisms and humans, many aspects remain contradictory, debatable or unclear due to the lack of complete high-quality networks. For example, does macromolecule connectivity in interactome maps reflect functional importance or sociological biases? Do biophysical interactions always reflect functional relationships? Here we propose to build a new, alternative view of an interactome map for functional importance or sociological biases? Do biophysical interactions reflect or unclear due to the lack of complete high-quality networks.

Motivation

Less than half of the yeast interactome covered

Binary PPI maps

Assembly from databases

Small-scale (min 2 supporting pieces of evidence)

Li-BM-13

Uetz et al., Nature 2000 (Uetz)

Tiro et al., PNAS 2001 (Tiro Core)

Yu et al., Science 2008 (Y1K)

Tarassov et al., Science 2008

Positive Reference Set (PRS)

116 interactions with ≥4 papers

Random Reference Set (RRS)

200 pairs identified from 1890 possible pairs from predicted yeast ORFeome

Coverage of binary PPI maps

Recovery rate of various binary PPI datasets

Large contribution from Li-BM-13 which is sociologically biased. All of the datasets are of good quality when compared to PRS/RRS. PPIs from PCA seem to be of a different nature.

Recovery of binary PPI in protein complexes and functional databases

Co-complex datasets


Collins et al., Molecular & Cellular Proteomics 2007: meta-analysis of two affinity purification/mass-spectrometory studies

Costanzo et al., Science 2010: Synthetic Genetic Array (SGA) and genetic interaction profiles based on PCC values

Functional dataset

High confidence list of functional interactions

PCC 0.3

≤ 60

≥ 100

Old edges

New edges

New nodes

Validation in an orthogonal assay: gPCA

1,286 PPI with confirmed ORF identity

Validation by G. pneumom Protein Complementation Assay (gPCA)

YI2 is a good quality dataset

YI2 contribution to YeRi:

nodes: 21% (467 out of 2,212)

edges: 32% (1,198 out of 3,712)

All yeast interactome datasets are biased for proteins that are more abundant and more studied. This suggest that part of these biases are biological.

Future plans

- Validate the YI2 dataset with another orthogonal assay to further asses the quality of our dataset

- Expand the screens to the complete search space for S. cerevisiae, including candidate proto-genies (~1,139 genes) (Carvunis et al., Nature 2012)

- Perform additional screens of the complete search space to generate a first Yeast Reference Interactome map (YeRi)

This work was supported by the grant R01HG006031