Identifying Conserved Protein Complexes between Species by Constructing Interolog Networks

InCoB 2013, Taicang China September 2013

Sriganesh SRIHARI

Institute for Molecular Bioscience, The University of Queensland, QLD, Australia

In collaboration with Phi Vu Nguyen and Hon Wai Leong

Department of Computer Science, National University of Singapore



THE UNIVERSITY OF QUEENSLAND AUSTRALIA





understand the underlying cellular machinery and organization.

Reconstructing the 'complexosome' still a long way to go!

- Yeast (most complete, most studied) 60-75%
 - Mainly missing are the membrane complexes
 - Wodak CYC 2008 (Pu et al., 2009), MIPS (Mewes et al., 2004)
- Human: 30-40%
- CORUM (Ruepp et al, 2011), Human soluble (Havugimana et al., 2012)
- Many complexes are conserved
 - Complexes are functional units
 - Useful to integrate evolutionary conservation to detect complexes

Interolog networks

- Integrating evolutionary information with PPI networks
 - Detect evolutionarily conserved protein interactions
 - Detect evolutionarily conserved complexes

Identifying conserved complexes between human and yeast

CONSTRUCTING INTEROLOG NETWORKS







Constructing Interolog Network



*Mainly sequence similarity used in the literature









Advantages of Using Domain Information for Interolog Network Construction









Improvement Over Earlier Orthology-network Methods including Sharan et al. (2006)

- Improved interolog network construction
 - Uses domain information apart from sequence similarity
 - Preserves many-to-many orthology relationships

Uses 'state-of-the-art' PPI network clustering algorithms

- CMC (Liu et al., Bioinformatics 2009)
- □ HACO (Wang et al., Cell Mol Proteomics 2009)
- □ MCL (van Dongen 2000/2004) and

MCL-CAw (Srihari et al., BMC Bioinformatics 2010)

Shown to perform significantly better than traditional clustering methods (Srihari et al., 2010, 2012, 2013)

Identifying conserved complexes between yeast and human

EXPERIMENTAL EVALUATION

PPI Datasets

Yeast:

#proteins: 5239

#interactions: 71636

Source: IntAct, BioGrid (Kerrien et al. 2007, Stark et al. 2011)

Database	# proteins	# interactions
IntAct (version Nov 13, 2012)	5276	18834
Biogrid (version 3.2.95, Nov 30, 2012)	5886	73923
IntAct ∪Biogrid	6332	83777
IntAct Biogrid	4620	8930
ICDScore(IntAct \cup Biogrid)	5239	71636

Human:

#proteins: 9764
#interactions: 192053
Source: BioGrid, HPRD
(Stark et al. 2011,
Prasad et al. 2009)

Database	# proteins	#interactions	
HPRD (Release 9, 2010)	9617	39184	
Biogrid (April 25, 2012)	12515	59027	
HPRD∪Biogrid	13624	76719	
HPRD∩Biogrid	8615	21491	
ICDScore(HPRD∪Biogrid)	8521	61868	
ICDEnrich(HPRD∪Biogrid)	9764	192053	

Protein Benchmark Complexes Datasets

- Wodak CYC2008 yeast complexes
 - 149 with size>3 (36.5%)
 - Total: 408
 - Pu S et al., NAR 2009

- CORUM human complexes
 - **722** with size>3 (39.1%)
 - Total: 1843
 - Ruepp et al. NAR 2008

COCIN Identifies More Conserved Complexes than Direct Clustering



COCIN Identifies More Conserved Complexes than Direct Clustering

Method	# Predicted	# Matched	Precision	# Gold standard	# Detected	Recall (of
	complexes	predictions		conserved	conserved	conserved
				complexes	complexes	complexes)
COCIN	71	36	50.7%	118	78	66.1%
СМС	1389	156	11.2%	118	66	55.9%
НАСО	1290	80	6.2%	118	36	30.5%
MCL-CAw/MCL	631	45	7.1%	118	24	20.3%

Similar results comparing against HACO and MCL-CAw/MCL

(Using Ensembl)











Cellular processes have evolved multi-fold from yeast to human.

- 1. Proteins have fused as well as segregated
- 2. Multiple proteins "invented" for buffering purposes
- 3. Co-functional proteins have parted ways (belong to different complexes)

Key relationships have been broken, new ones formed					
(e.g. cell cycle and DNA damage repair)	Human				
Conserved by sequence similarity	v				
Conservation of domains	s				

THANK YOU

National University of Singapore:

Phi Vu Nguyen Prof Hon Wai Leong

Thank You...

Acknowledgements

Institute for Molecular Bioscience

Prof Mark Ragan and his group



Queensland Institute of Medical Research

Australian Government

Medical Research Council

UQ Centre for Clinical Research Dr Peter T. Simpson

Queensland Institute of Medical Research Prof Kum Kum Khanna and her group

NHMRC grant to Dr Peter T. Simpson & Prof Mark A. Ragan

Thank You...

