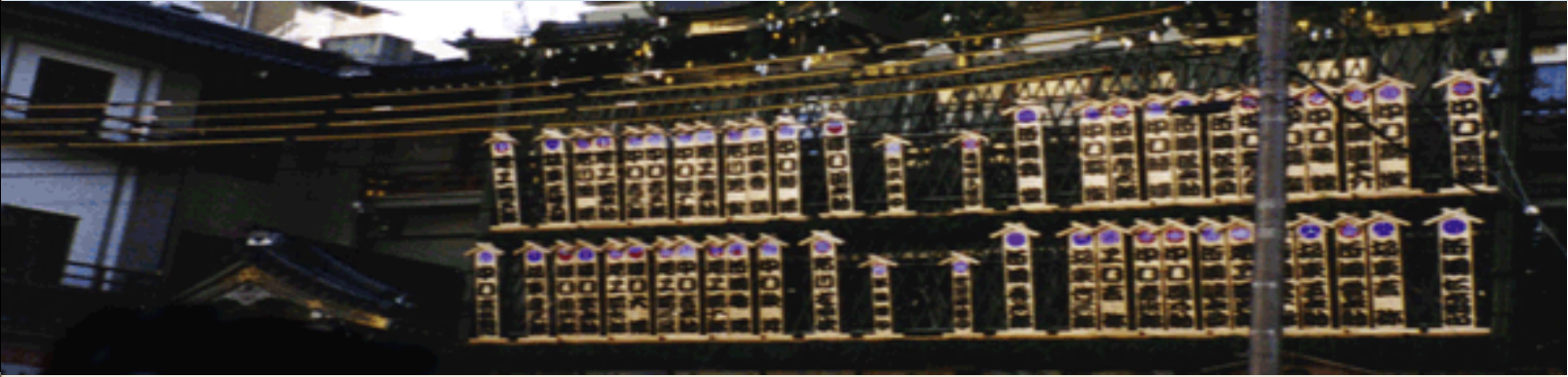


redundancy, control and collective computation



in network dynamics

**luis m. rocha**

school of informatics & computing

indiana university, bloomington, usa

and

instituto gulbenkian de ciência

oeiras, portugal

**Informatics**  
luis rocha 2014



INDIANA  
UNIVERSITY

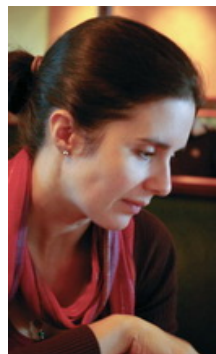


[rocha@indiana.edu](mailto:rocha@indiana.edu)  
<http://informatics.indiana.edu/rocha>

CASCI group



Manuel Marques-Pita



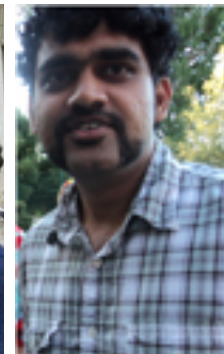
Anália Lourenço



Tiago Simas



Artemy Kolchinsky



Santosh Manicka



Alexander Gates



Jason Yoder



Ian Wood



Rion Brattig Correia



Yizhi Jing



Thomas Parmer



PERSISTENT



Fundação para a Ciência e a Tecnologia  
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR



FUNDAÇÃO CALOUSTE GULBENKIAN

Indiana University Collaborative Research Grants

[rocha@indiana.edu](mailto:rocha@indiana.edu)

<http://informatics.indiana.edu/rocha>

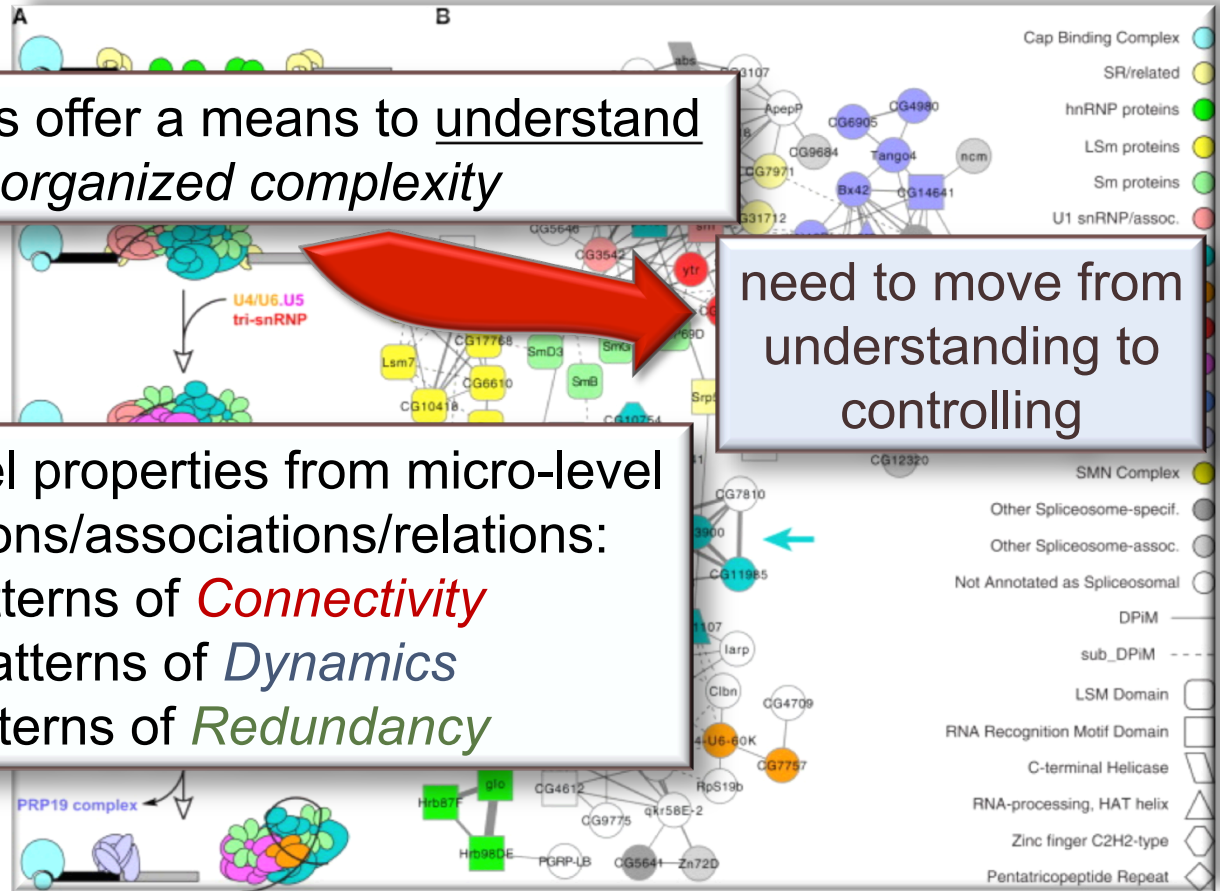
Informatics  
luis rocha 2014

INDIANA  
UNIVERSITY



large-scale *drosophila* protein interaction Map (DPiM)

Guruharsha et al [2011]. "A Protein Complex Network of *Drosophila melanogaster*." *Cell*. **147**(3):690-703.



Networks offer a means to understand *organized complexity*

need to move from understanding to controlling

Macro-level properties from micro-level interactions/associations/relations:  
 Patterns of *Connectivity*  
 Patterns of *Dynamics*  
 Patterns of *Redundancy*

Modularity of the spliceosome subnetwork: 12 well-connected clusters representing interaction of snRNPs with pre-mRNA and other proteins in the process of splicing introns.



**Informatics**  
luis rocha 2014



# PATTERNS OF REDUNDANCY IN DYNAMICS



# the drosophila segment polarity network

## an automata network model built from qualitative data



Based on the ODE model of von Dassow et al. (2000), consists of 4-cell parasegments, each cell with 15 interacting genes and proteins.

**2<sup>60</sup> network configurations**

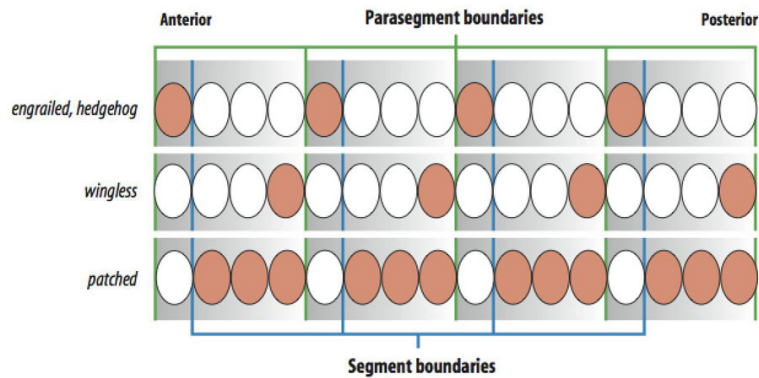
Reproduces wild-type and mutant gene expression patterns in development of fruit fly

2 intercellular inputs: **nhh** (*hedgehog*), **nWG** (*wingless*)

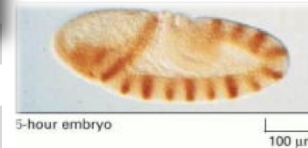
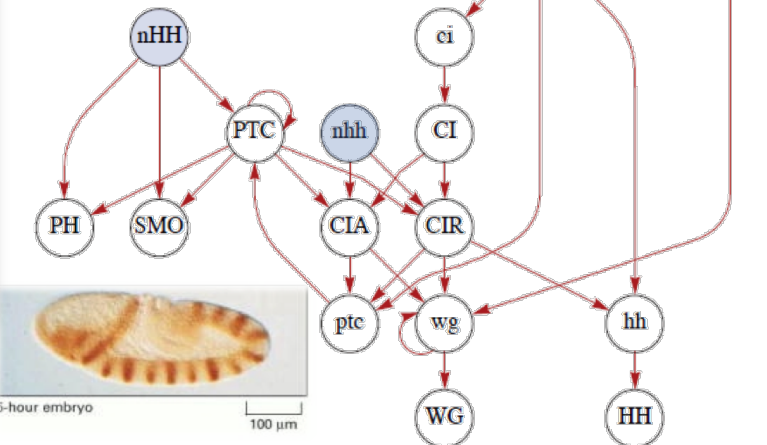
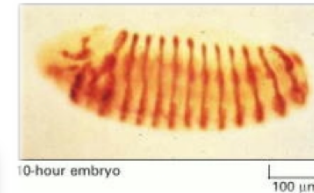
1 intracellular input: **SLP** (*sloppy paired*)



Node	State – TransitionFunction
$SLP_i^{t+1}$	$\leftarrow 0$ if $i=1 \vee i=2$ ; 1 if $i=3 \vee i=4$ ;
$wg_i^{t+1}$	$\leftarrow (CIA_i^t \wedge SLP_i^t \wedge \neg CIR_i^t) \vee (wg_i^t \wedge (CIA_i^t \vee SLP_i^t) \wedge \neg CIR_i^t)$



$Ci_i^{t+1}$	$\leftarrow ci_i^t$
$CIA_i^{t+1}$	$\leftarrow CIA_i^t \wedge (\neg PTC_i^t \vee hh_{i-1}^t \vee hh_{i+1}^t \vee HH_{i-1}^t \vee HH_{i+1}^t)$
$CIR_i^{t+1}$	$\leftarrow CIR_i^t \wedge PTC_i^t \wedge \neg hh_{i-1}^t \wedge \neg hh_{i+1}^t \wedge \neg HH_{i-1}^t \wedge \neg HH_{i+1}^t$



# quantifying micro-level canalization

input redundancy, effective connectivity and input symmetry

**A** Look-Up Table

$f_1$	0	0	1	0	0	1	: 1
$f_2$	0	0	1	0	1	0	: 1
$f_3$	0	0	1	0	1	1	: 1
$f_4$	0	0	1	1	0	0	: 1
$f_5$	0	0	1	1	0	1	: 1
$f_6$	0	0	1	1	1	0	: 1
$f_7$	0	0	1	1	1	1	: 1
$f_8$	1	0	1	0	0	0	: 1
$f_9$	1	0	1	0	0	1	: 1
$f_{10}$	1	0	1	0	1	0	: 1
$f_{11}$	1	0	1	0	1	1	: 1
$f_{12}$	1	0	1	1	0	0	: 1
$f_{13}$	1	0	1	1	0	1	: 1
$f_{14}$	1	0	1	1	1	0	: 1

$\Theta_1$   $\Upsilon_9$

**B** Wildcard Schemata

$f'_1$	#	0	1	#	0	1	: 1
$f'_2$	#	0	1	#	1	0	: 1
$f'_3$	#	0	1	0	#	1	: 1
$f'_4$	#	0	1	0	1	#	: 1
$f'_5$	#	0	1	1	#	0	: 1
$f'_6$	#	0	1	1	0	#	: 1
$f'_7$	0	0	1	#	#	1	: 1
$f'_8$	0	0	1	#	1	#	: 1
$f'_9$	0	0	1	1	#	#	: 1
$f'_{10}$	1	0	1	#	#	0	: 1
$f'_{11}$	1	0	1	#	0	#	: 1
$f'_{12}$	1	0	1	0	#	#	: 1

$F'$

$$k(x) = 6$$

$$k_r(x) = \frac{\sum_{f_\alpha \in F} \max_{\theta | f_\alpha \in \Theta_\theta} (n_\theta^\#)}{2^k}$$

$$k_e(x) = k(x) - k_r(x)$$

$$k_s(x) = \frac{\sum_{f_\alpha \in F} \min_{\theta | f_\alpha \in \Theta_\theta} |n_\theta^o|}{2^k}$$

**C** Wildcard & Position Free Schemata

$f''_1$	#	0	1	#	0	1	: 1
---------	---	---	---	---	---	---	-----

$F''$

- Measuring two forms of canalization
  - $K_r = 2$
  - $K_e = 6 - 2 = 4$
  - $K_s = 4$

Prime Implicants (Quine-McCluskey) plus group invariance

Informatics  
luis rocha 2014

In biological Boolean network models

- extracting micro-level canalization
  - drosophila segment polarity genes network

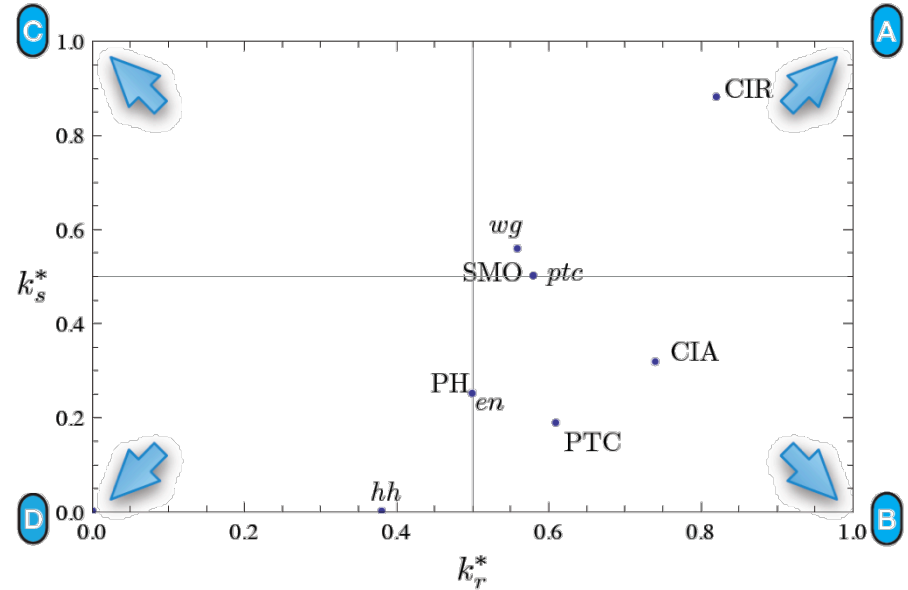
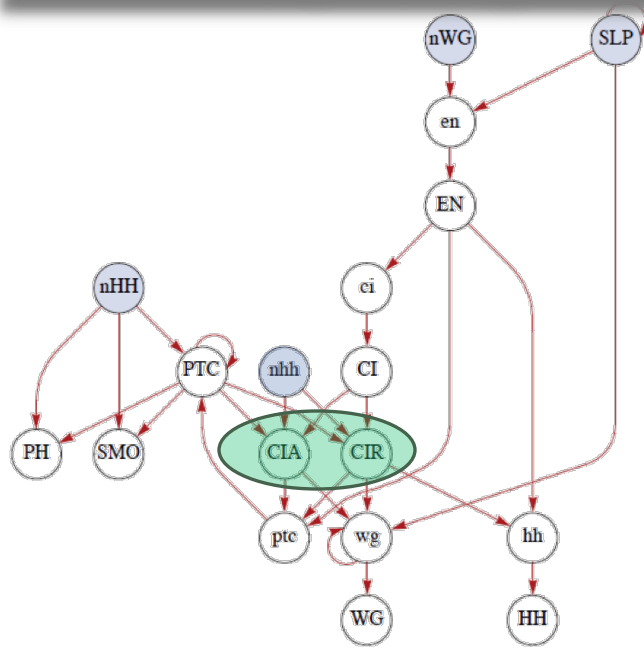
node	inhibition	expression	$k$	$k_e$	$k_r$	$k_s$	$k_r^*$	$k_s^*$
<i>wg</i>	$f''_{2:1}$ $f''_{2:2}$	$f''_{2:3}$	4	1.75	2.25	2.25	0.56	0.56
PTC	$f''_{9:1}$ $f''_{9:2}$	$f''_{9:3}$ $f''_{9:4}$	4	1.56	2.44	0.75	0.61	0.19
PH	$f''_{10:1}$ $f''_{10:2}$	$f''_{10:3}$	3	1.5	1.5	0.75	0.5	0.25
SMO	$f''_{11:1}$	$f''_{11:2}$ $f''_{11:3}$	3	1.25	1.75	1.5	0.58	0.5
<i>ci</i>	$f''_{12:1}$	$f''_{12:2}$	1	1	0	0	0	0
CI	$f''_{13:1}$	$f''_{13:2}$	1	1	0	0	0	0
CIA	$f''_{14:1}$ $f''_{14:2}$	$f''_{14:3}$ $f''_{14:4}$	6	1.55	4.45	1.875	0.74	0.32
CIR	$f''_{15:1}$ $f''_{15:2}$	$f''_{15:3}$	6	1.08	4.92	5.25	0.82	0.88



# per-node schema redescription

In biological Boolean network models

- extracting micro-level canalization
  - drosophila segment polarity genes network

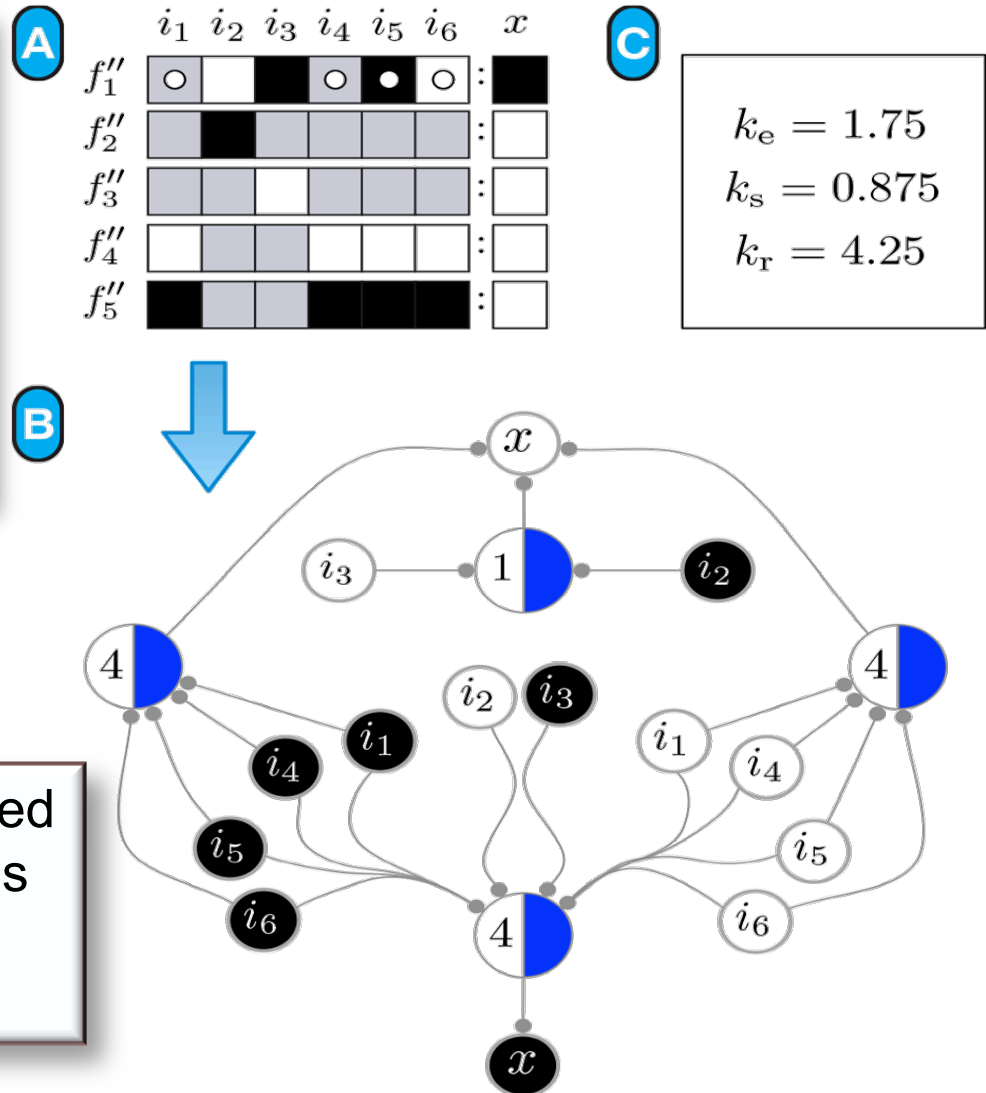


node	inhibition	expression	$k$	$k_e$	$k_r$	$k_s$	$k_r^*$	$k_s^*$																																																								
CIA	$f''_{14:1}$ <table border="1"> <tr> <td>PTC</td><td>CI</td><td>hh<sub>i-1</sub></td><td>hh<sub>i+1</sub></td><td>HH<sub>i-1</sub></td><td>HH<sub>i+1</sub></td><td>CIA</td> </tr> <tr> <td>█</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td> </tr> </table> $f''_{14:2}$ <table border="1"> <tr> <td>PTC</td><td>CI</td><td>hh<sub>i-1</sub></td><td>hh<sub>i+1</sub></td><td>HH<sub>i-1</sub></td><td>HH<sub>i+1</sub></td><td>CIA</td> </tr> <tr> <td>█</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td> </tr> </table>	PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIA	█	█	█	█	█	█	█	PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIA	█	█	█	█	█	█	█	$f''_{14:3}$ <table border="1"> <tr> <td>PTC</td><td>CI</td><td>hh<sub>i-1</sub></td><td>hh<sub>i+1</sub></td><td>HH<sub>i-1</sub></td><td>HH<sub>i+1</sub></td><td>CIA</td> </tr> <tr> <td>█</td><td>█</td><td>●</td><td>○</td><td>○</td><td>○</td><td>█</td> </tr> </table> $f''_{14:4}$ <table border="1"> <tr> <td>PTC</td><td>CI</td><td>hh<sub>i-1</sub></td><td>hh<sub>i+1</sub></td><td>HH<sub>i-1</sub></td><td>HH<sub>i+1</sub></td><td>CIA</td> </tr> <tr> <td>█</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td> </tr> </table>	PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIA	█	█	●	○	○	○	█	PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIA	█	█	█	█	█	█	█	6	1.55	4.45	1.875	0.74	0.32
PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIA																																																										
█	█	█	█	█	█	█																																																										
PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIA																																																										
█	█	█	█	█	█	█																																																										
PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIA																																																										
█	█	●	○	○	○	█																																																										
PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIA																																																										
█	█	█	█	█	█	█																																																										
CIR	$f''_{15:1}$ <table border="1"> <tr> <td>PTC</td><td>CI</td><td>hh<sub>i-1</sub></td><td>hh<sub>i+1</sub></td><td>HH<sub>i-1</sub></td><td>HH<sub>i+1</sub></td><td>CIR</td> </tr> <tr> <td>█</td><td>█</td><td>●</td><td>○</td><td>○</td><td>○</td><td>█</td> </tr> </table> $f''_{15:2}$ <table border="1"> <tr> <td>PTC</td><td>CI</td><td>hh<sub>i-1</sub></td><td>hh<sub>i+1</sub></td><td>HH<sub>i-1</sub></td><td>HH<sub>i+1</sub></td><td>CIR</td> </tr> <tr> <td>○</td><td>○</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td> </tr> </table>	PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIR	█	█	●	○	○	○	█	PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIR	○	○	█	█	█	█	█	$f''_{15:3}$ <table border="1"> <tr> <td>PTC</td><td>CI</td><td>hh<sub>i-1</sub></td><td>hh<sub>i+1</sub></td><td>HH<sub>i-1</sub></td><td>HH<sub>i+1</sub></td><td>CIR</td> </tr> <tr> <td>█</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td><td>█</td> </tr> </table>	PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIR	█	█	█	█	█	█	█	6	1.08	4.92	5.25	0.82	0.88														
PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIR																																																										
█	█	●	○	○	○	█																																																										
PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIR																																																										
○	○	█	█	█	█	█																																																										
PTC	CI	hh <sub>i-1</sub>	hh <sub>i+1</sub>	HH <sub>i-1</sub>	HH <sub>i+1</sub>	CIR																																																										
█	█	█	█	█	█	█																																																										

# canalization map as minimal control

## two-symbol schemata as threshold networks

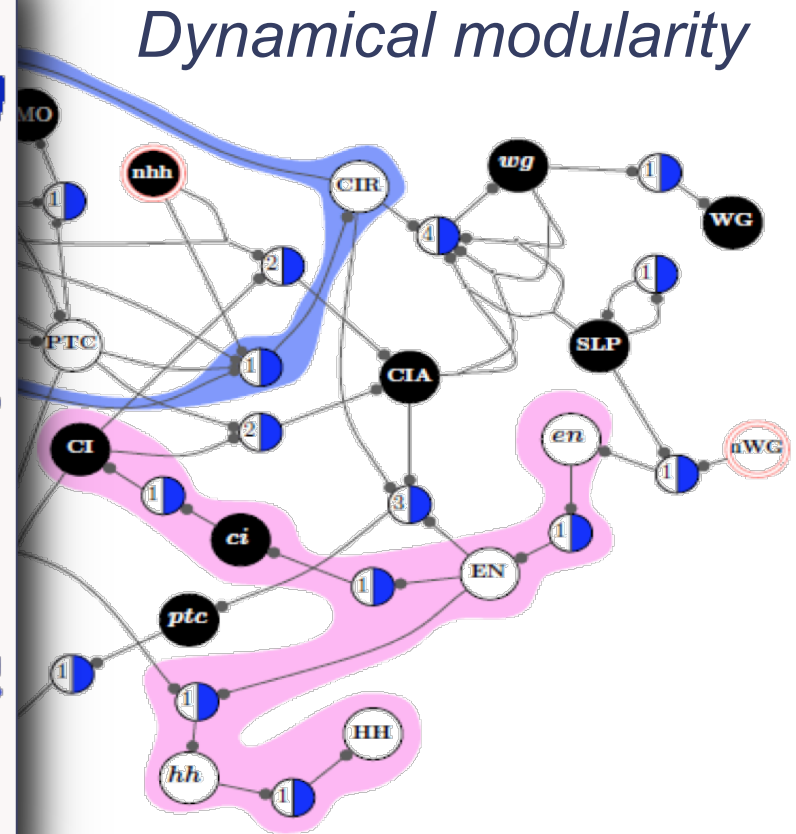
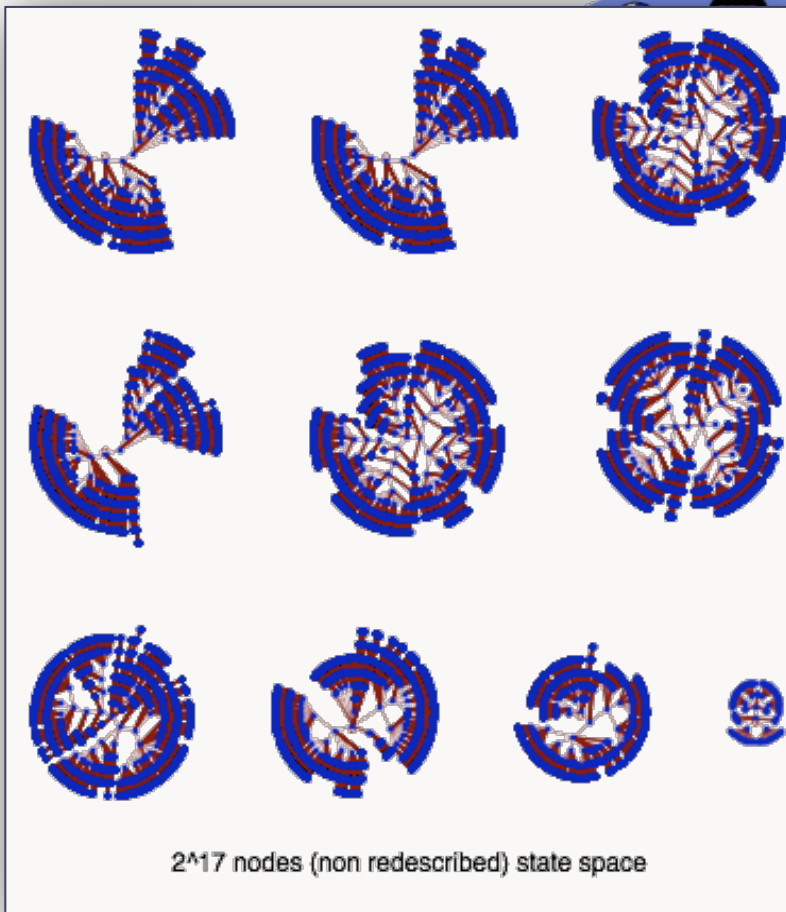
- understanding natural “computation”
  - How cells compute
  - minimal wiring (control) of micro-level



Each schema represented by conjunction of literals and symmetric group constraints

# (macro-level) dynamics canalization map from per-node schemata redescription

- Full dynamics (of single-cell model) captured by threshold network of  $2*N+M$  nodes



Informatics  
luis rocha 2014







**Informatics**  
luis rocha 2014

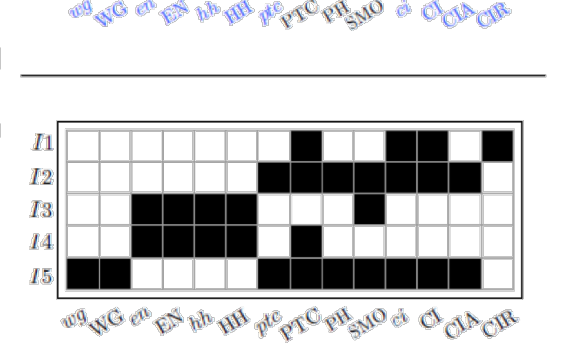
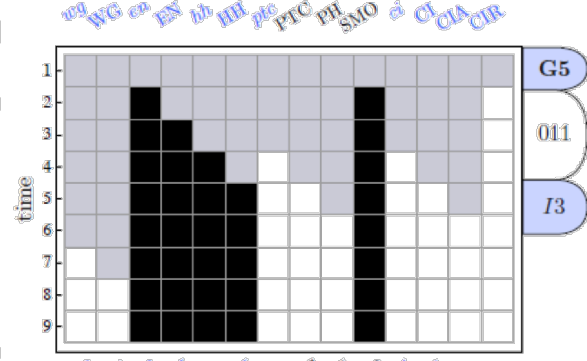
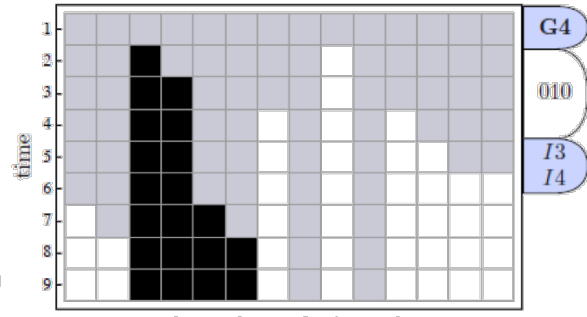
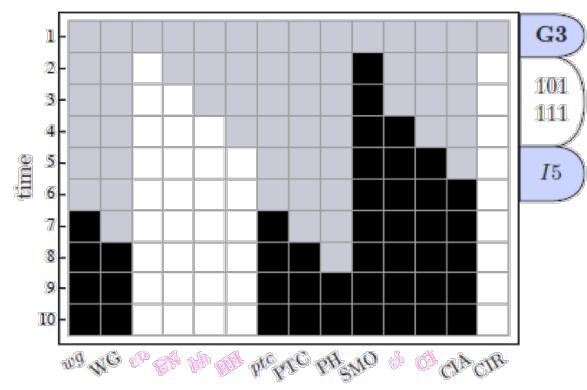
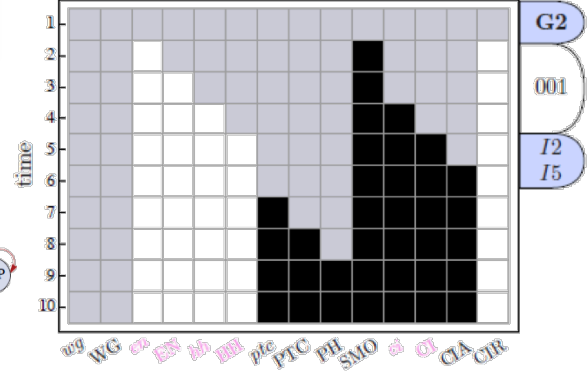
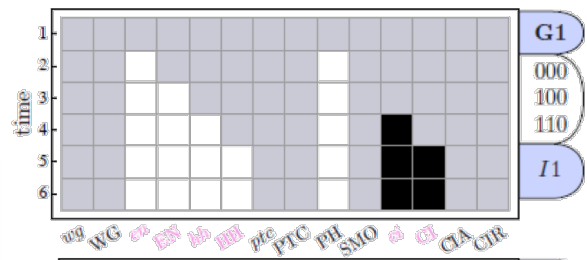
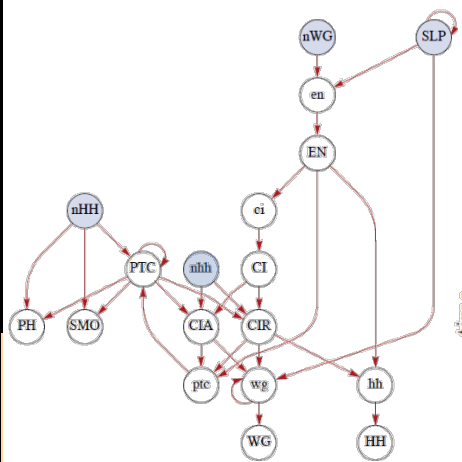


# CONTROL FROM PATTERNS OF REDUNDANCY IN DYNAMICS

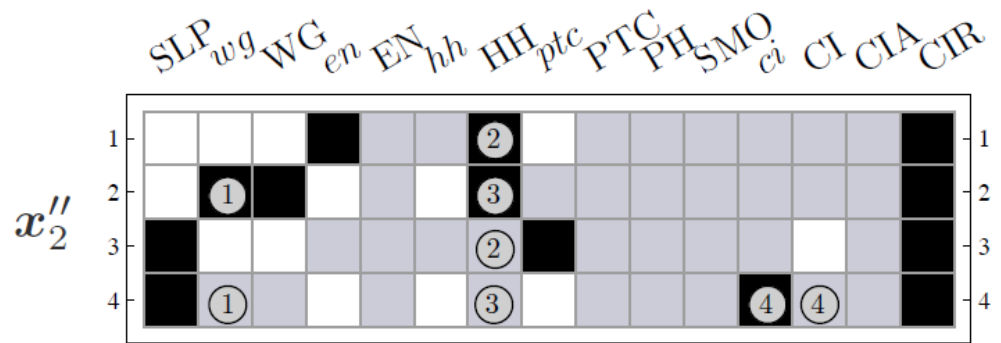
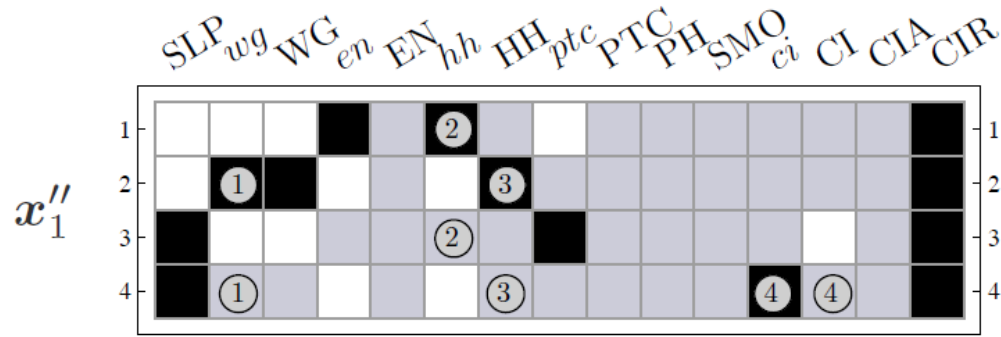
# Dynamical unfolding from partial information

- inputs in drosophila segment polarity net: SLP, nWG, nhh

How much control certain nodes have on network dynamics.



## minimal conditions (“pre-patterns”) for wild-type attractor



Less than half of the nodes are needed to ensure convergence to **wt**, which is very **robust** and **larger** than previously thought.

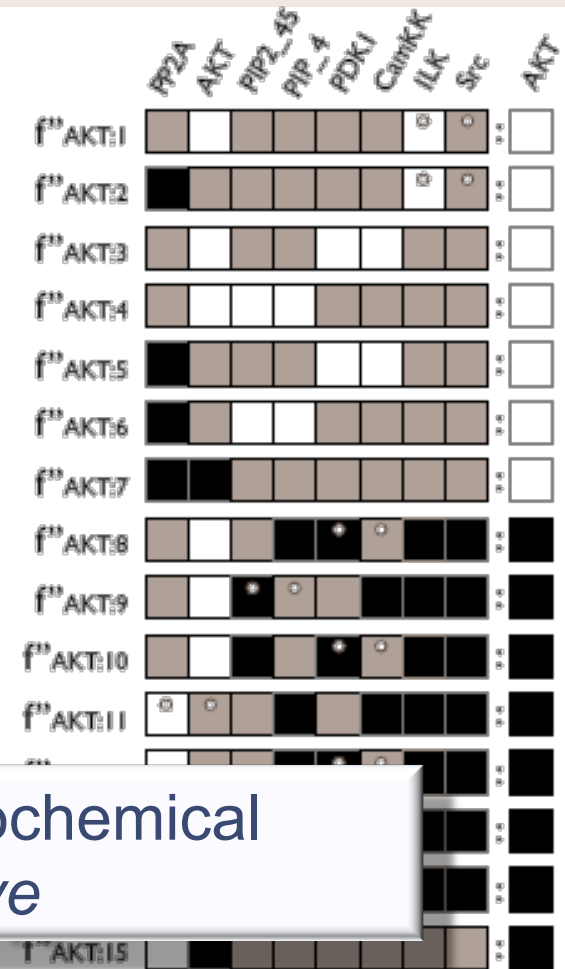
But losing one essential input (enputs) leads to a “unspecific” attractor. In this case could go to wt, wt (ptc mutant) or no segmentation.



# redundancy in intracellular signaling networks

## canalization

- Activation of AKT in generic fibroblasts (130 node BN)
  - LUT of  $2^8=256$  entries redescrbed by only 15 schemata
    - Large amount of **canalization**
  - Very few actual inputs need to be known to determine state-transition



**Upcoming work:** analysis of biochemical models in the entire *cell collective*

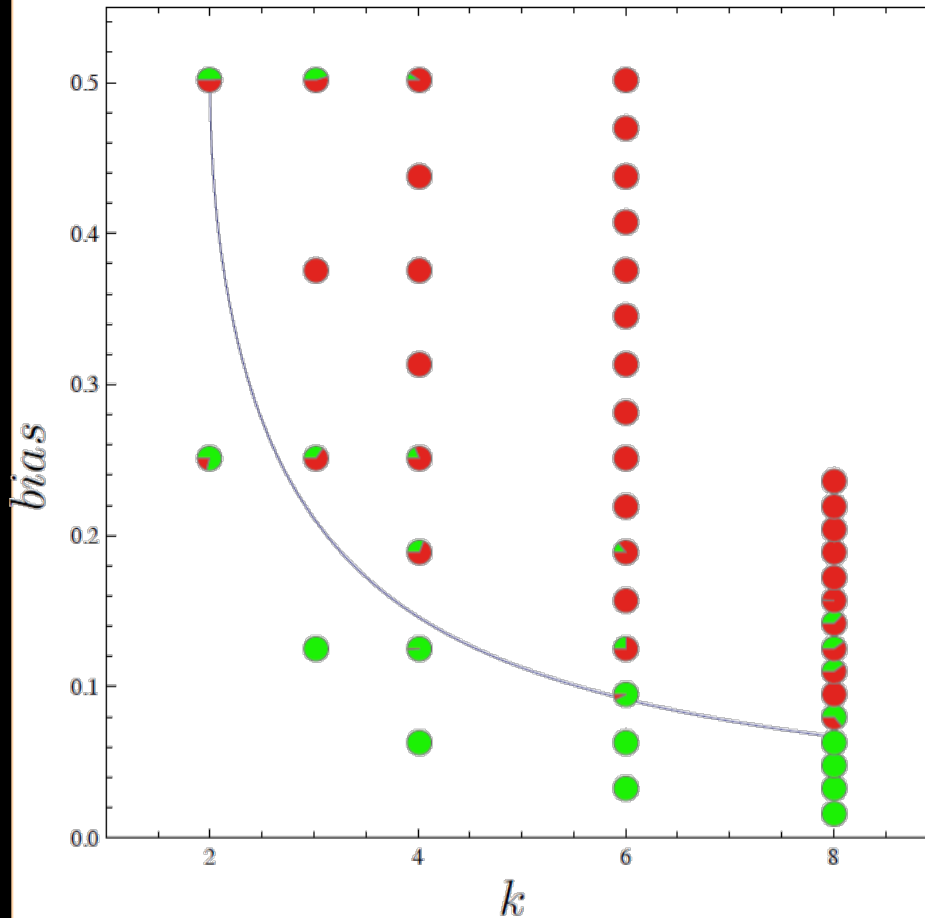


**Informatics**  
luis rocha 2014



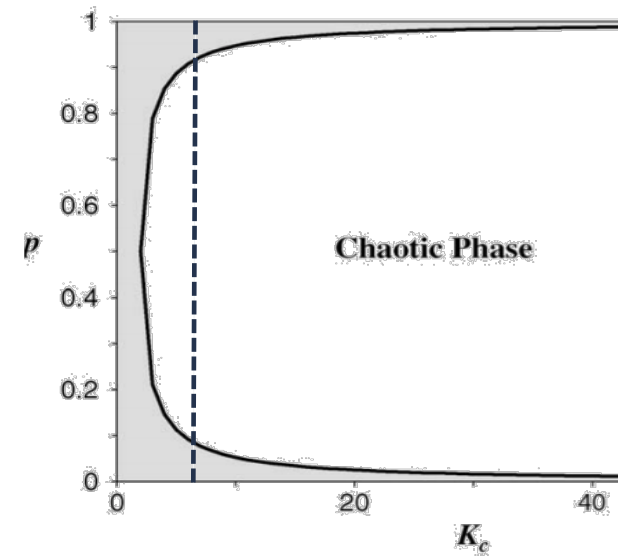
# PATTERNS OF DYNAMICS

## Current theory



$$p = \frac{1}{2} \left( 1 - \sqrt{1 - \frac{2}{k}} \right)$$

Aldana, M. [2003]. *Physica D.* **185**: 45–66



Informatics  
luis rocha 2014





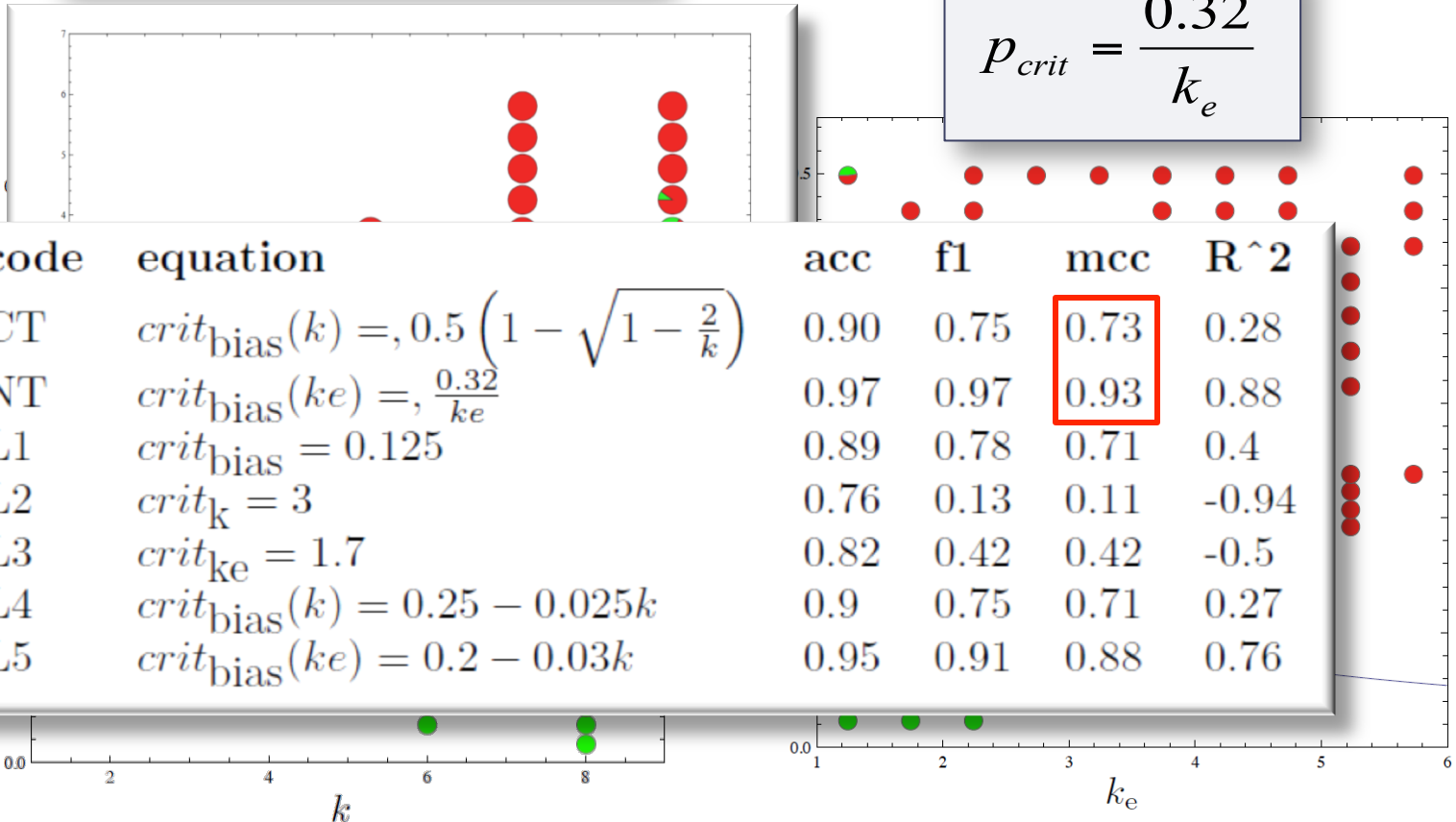
# criticality in the presence of canalization

input redundancy, effective connectivity

$$k_r(x) = \frac{\sum_{f_\alpha \in F} \max_{\theta | f_\alpha \in \Theta_\theta} (n_\theta^\#)}{2^k}$$

$$k_e(x) = k(x) - k_r(x)$$

$$p_{crit} = \frac{0.32}{k_e}$$



Informatics  
luis rocha 2014



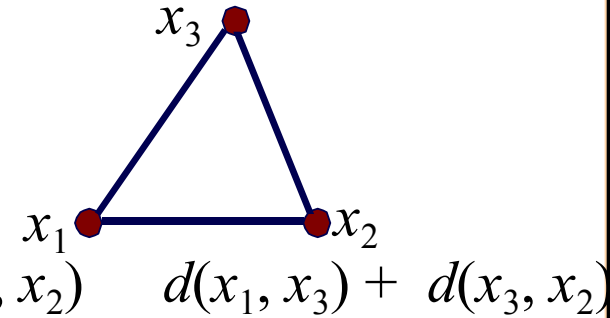
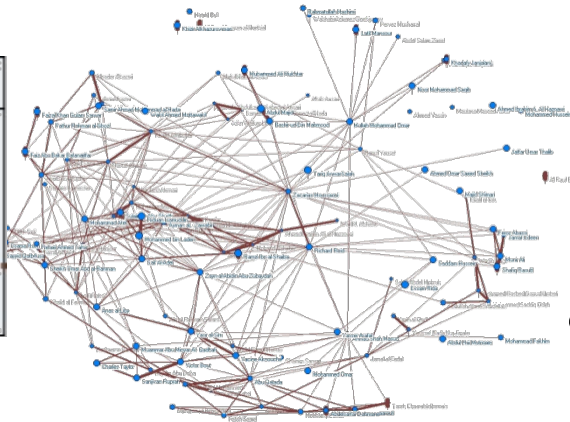
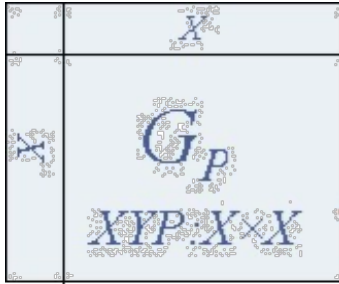
**Informatics**  
luis rocha 2014



# PATTERNS OF CONNECTIVITY

from weighted graphs

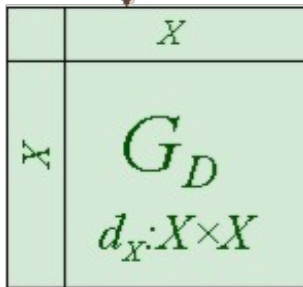
Proximity



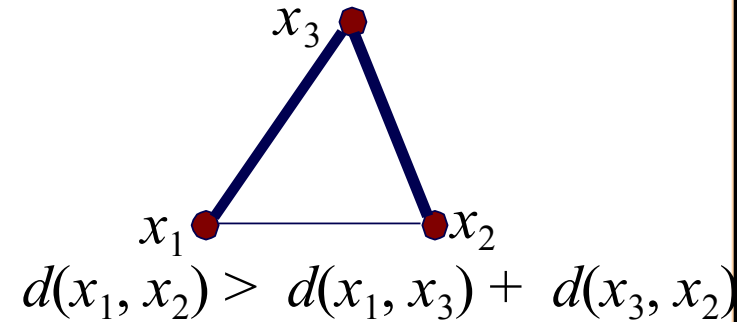
Metric

$$\phi : d_X(x_i, x_j) = \frac{1}{XYP(x_i, x_j)} - 1$$

$\phi$  is a nonlinear **distance function**: nonnegative, symmetric, antireflexive ( $d(x, x) = 0$ )



Distance



Semi-metric

Rocha & Bollen [2001] In: *SFI Series*. Segel & Cohen (Eds.), 305-334.

Rocha, L.M. [2002]. In: *Soft Computing Agents*: 137-163.

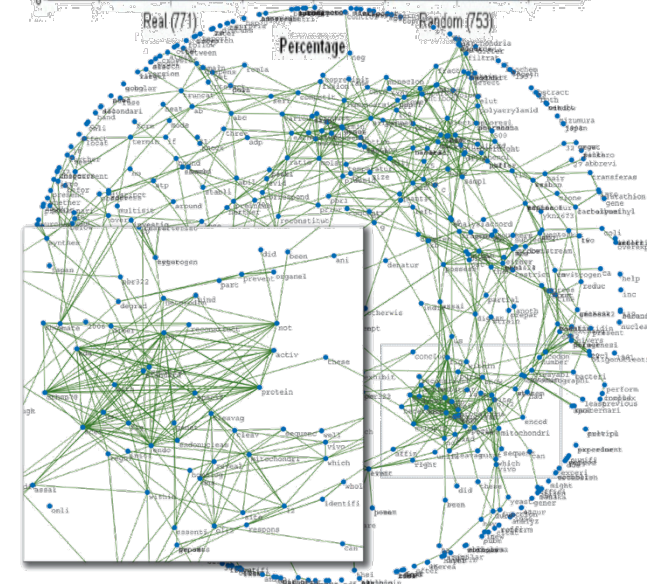
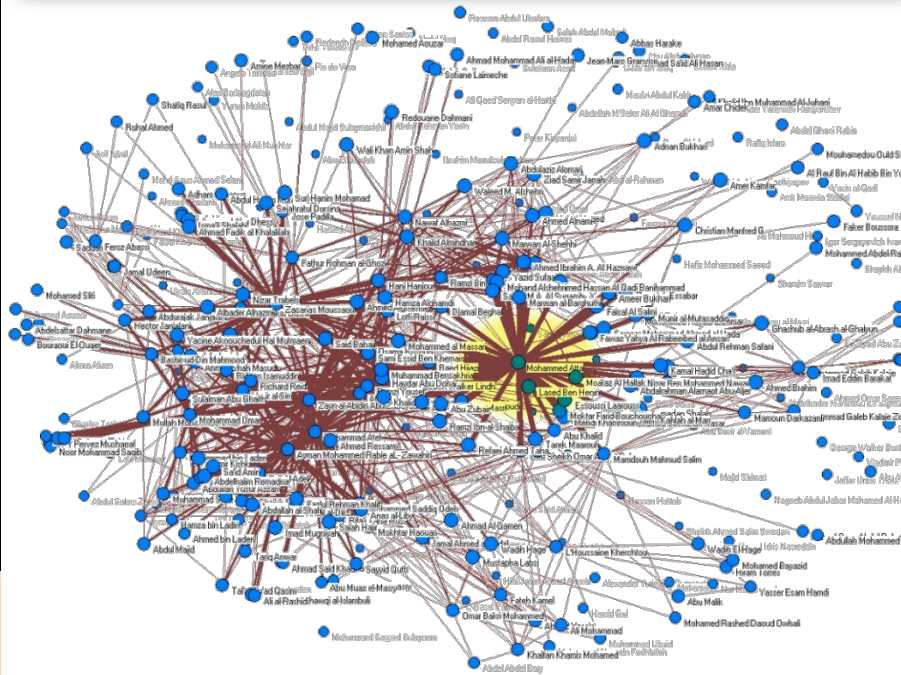
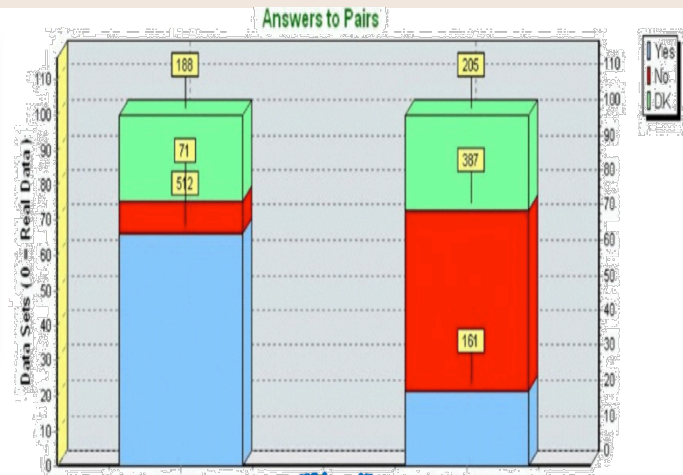
Rocha, L.M. et al [2005]. *IEEE Web Intelligence (WI'05)*: 565-571.





## latent associations in data

- indirectly connected items
  - Higher change of future strength
- Applications
  - Recommender systems
  - Social biochemical networks

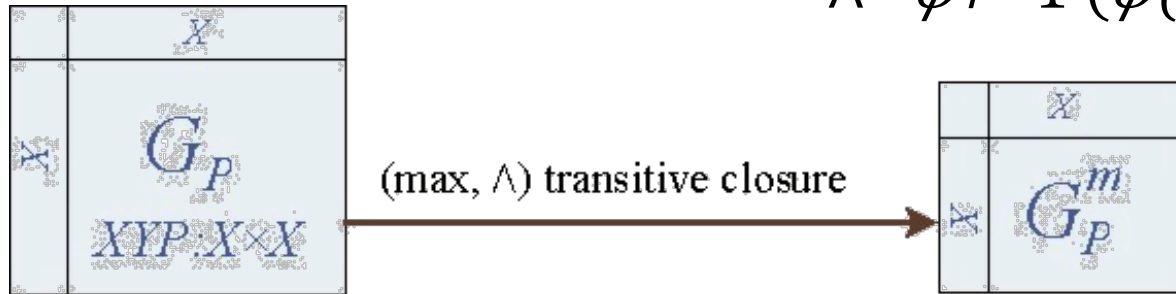


Rocha, L.M. et al [2005]. *IEEE Web Intelligence (WI'05)*: 565-571.  
 Simas & Rocha [2012]. *IEEE Web Intelligence (WI'12)*. 175-179.  
 Abi-Haidar, A et al. [2008] *Genome Biology* 9(Suppl 2):S11.

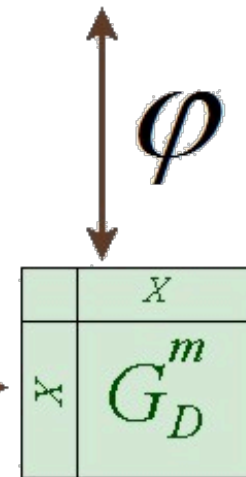
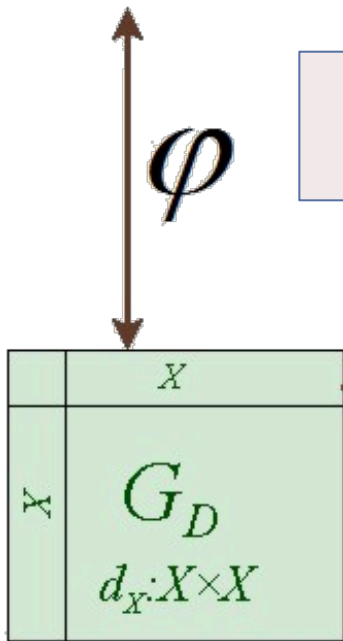
shortest-path closures ( $\vee$ =maximum)

generalized distance closure with APSP (Dijkstra)

$$\wedge = \varphi \uparrow - 1 (\varphi(a) + \varphi(b))$$



explores all possible *path length* measures



(min, +) metric closure

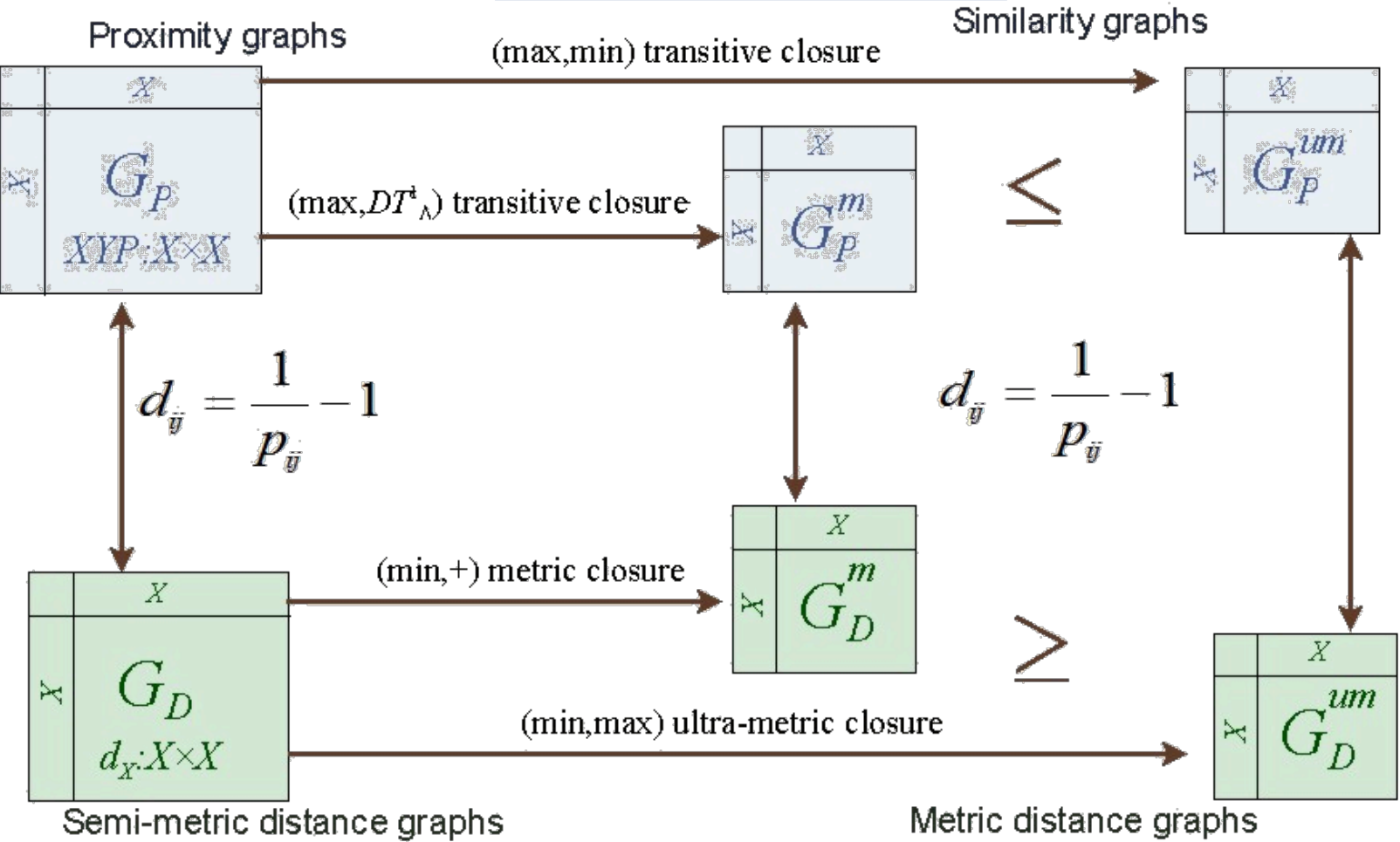
transitive closure converges in finite time

Semi-metric distance graphs

Informatics  
luis rocha 2014

most aggressive

shortest-path where *path length* is weakest link



Informatics  
luis rocha 2014





**Informatics**  
luis rocha 2014



# PATTERNS OF REDUNDANCY

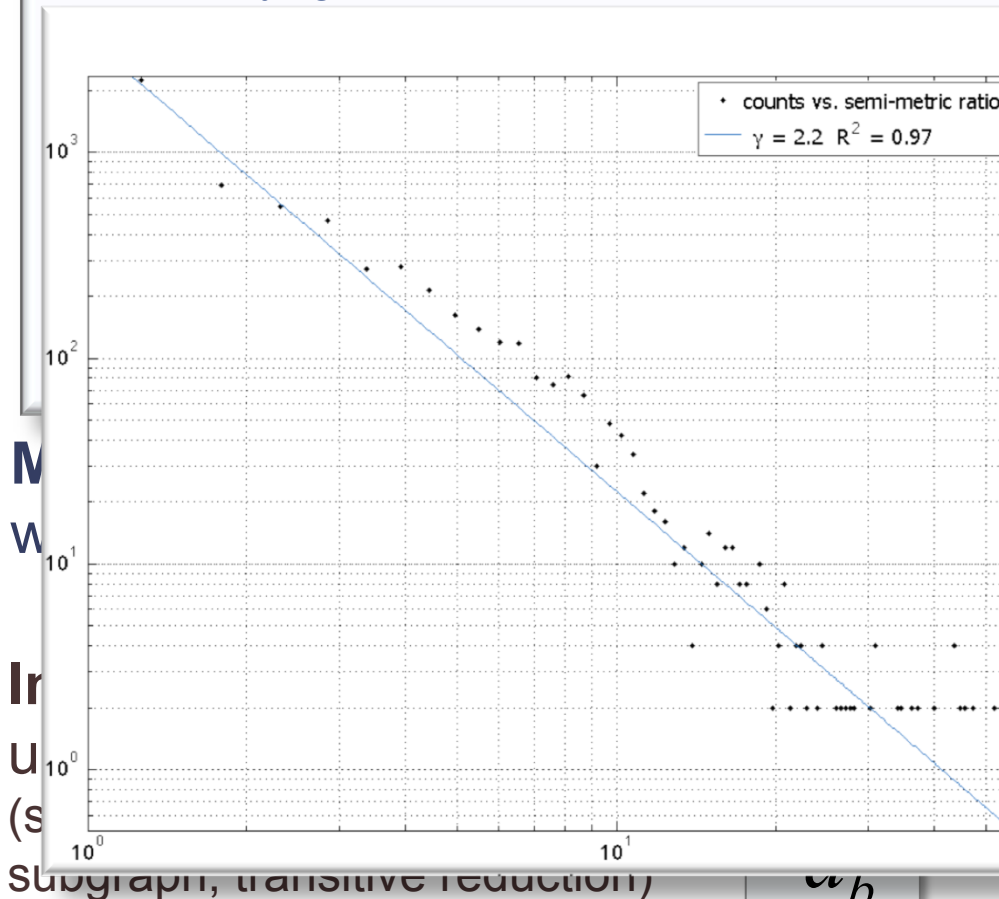


# semi-metric (semi-triangular) behavior

## and structure of weighted complex networks

### ■ Semi-metric (semi-triangular) edges:

- Redundant for shortest-path computation (distance closure)
- Null edge betweenness centrality
- Varying *semi-metric distortion*



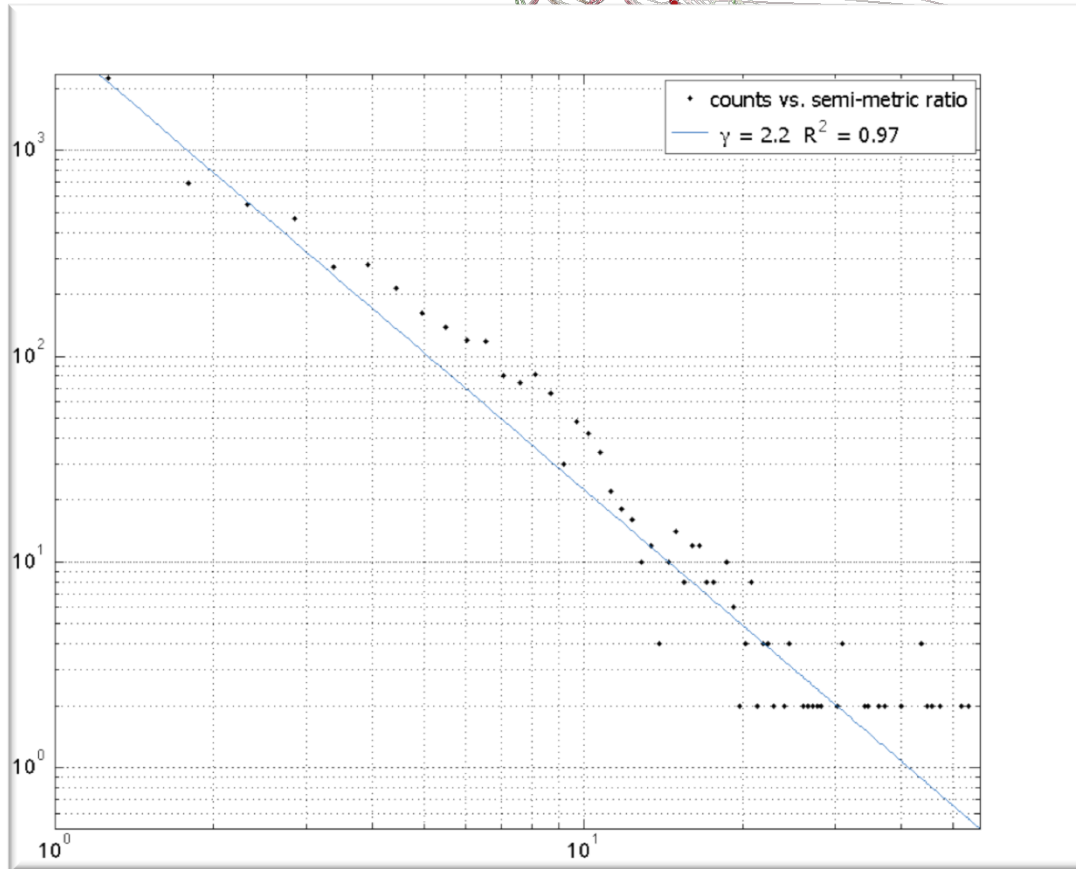
$$s \downarrow i, j = d \downarrow i, j$$

b-graph for

$$d^{mc}$$

US airport Network

Colizza,, Pastor-Satorras,, Vespignani [2007].  
*Nature Physics* 3, 276-282.;



Available airplane seats between US cities

5% (83%) of edges are semi-metric (semi-ultra-metric) and removed from backbone

Simas, T. [2012]. *PhD Thesis*. Indiana University.

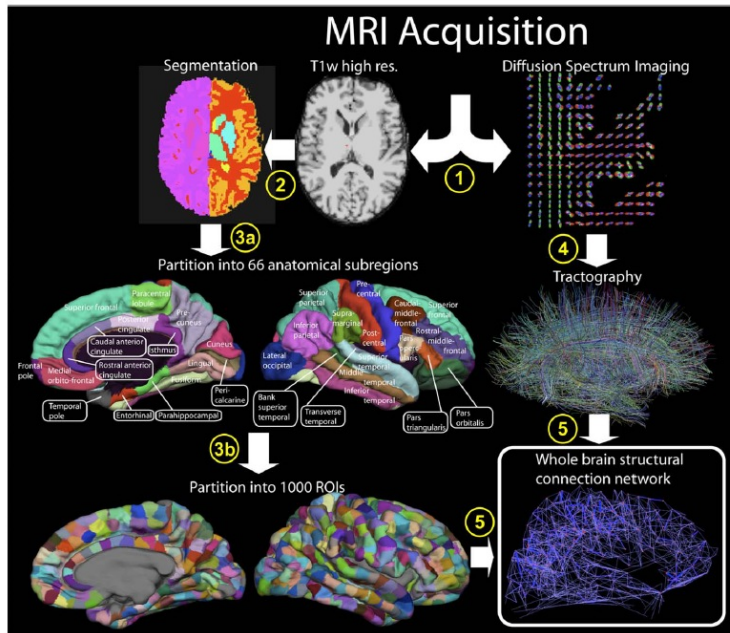
Simas, Ciampaglia, Sporns & Rocha [2014]. In Preparation.



Informatics  
luis rocha 2014

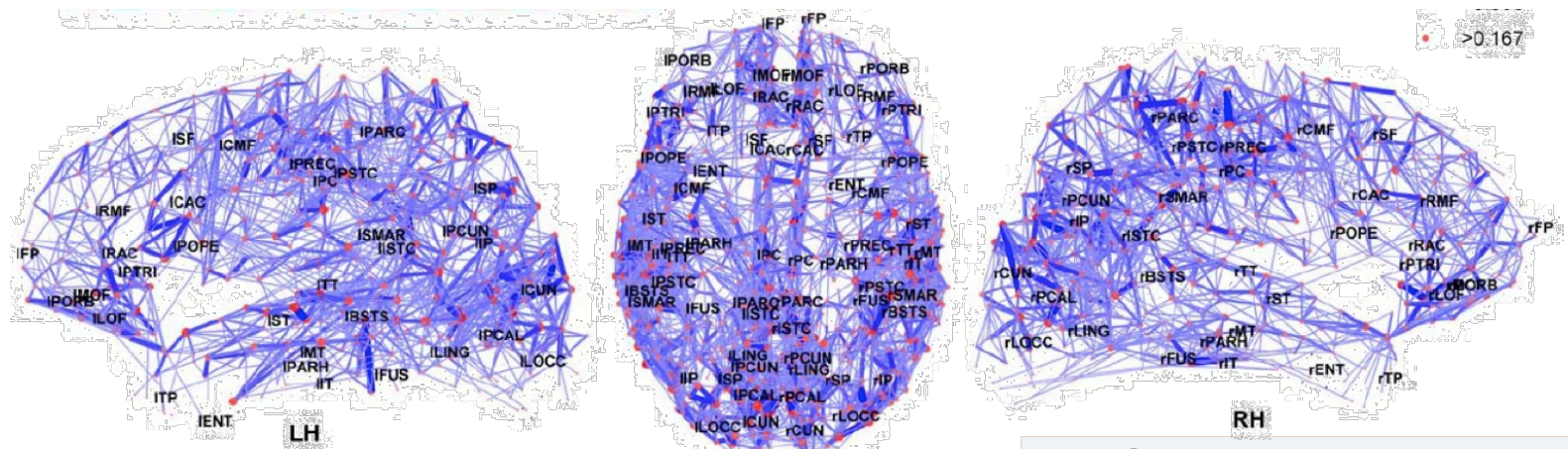


## Full network



cortico-cortical  
axonal pathways from  
diffusion spectrum  
Imaging (DSI)

Nodes: functionally  
specialized regions



Hagmann et al. [2008]. *PLoS Biol* 6(7): e159.

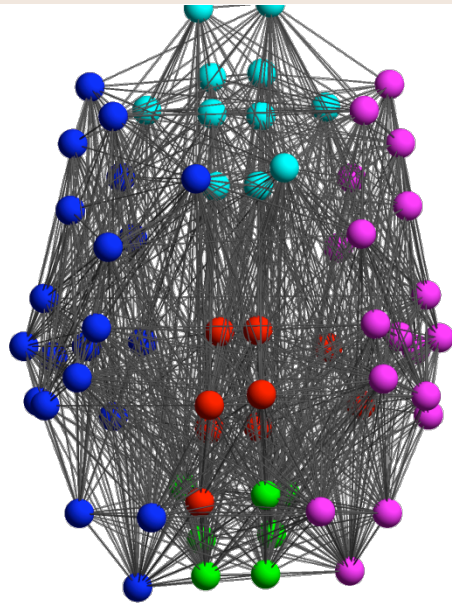


Informatics  
luis rocha 2014



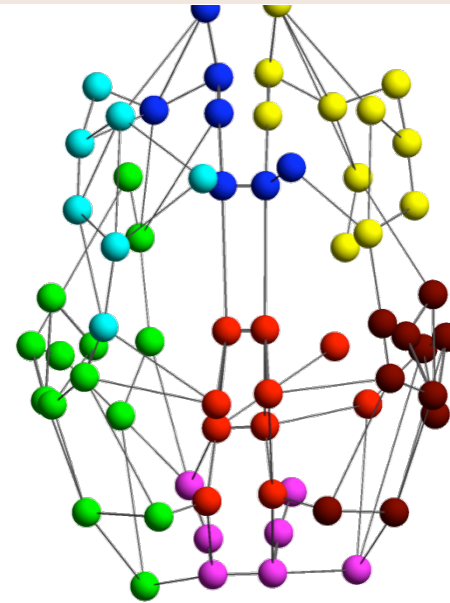


Human Cerebral Cortex Network



**Original**

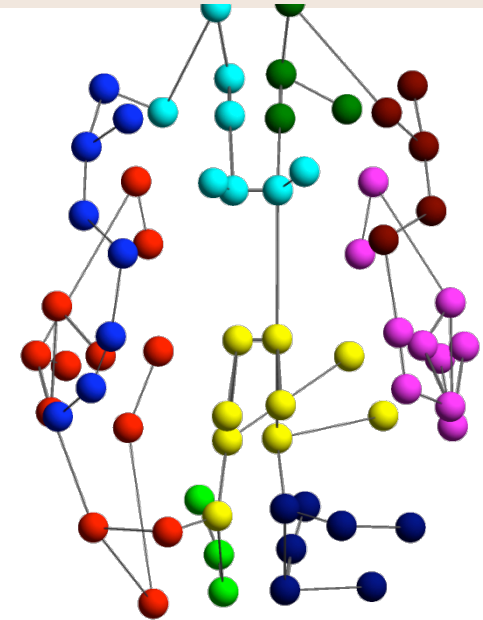
cortico-cortical  
axonal pathways from  
diffusion spectrum  
Imaging (DSI)



**Metric Backbone**

91% of edges are semi-metric  
(removed from backbone)!!!

94% of edges are semi-ultra-metric



**Ultra-Metric  
Backbone**

Simas, T. [2012]. *PhD Thesis*. Indiana University.

Simas, Ciampaglia, Sporns & Rocha [2014]. In Preparation.

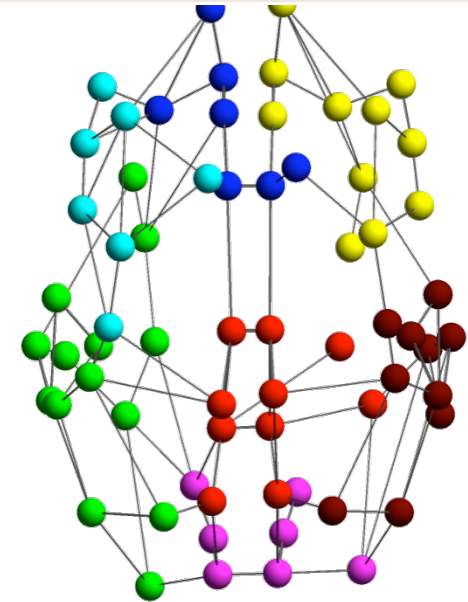
[rocha@indiana.edu](mailto:rocha@indiana.edu)

<http://informatics.indiana.edu/rocha>



## Semi-metricity and modularity

#id	Network	# Nodes	SM	USM
1	USN	500	75%	83%
2	HCN	66	91%	94%
3	HBFN	116	85%	98%
4	C-Elegans	297	31%	45%
5	BKF	58	85%	91%
6	ARP-IPP	1,702	71%	88%
7	ARP-PIP	382	73%	93%
8	ARP-Keywords	500	96%	99%
9	WordNet	150	85%	93%
10	SCN	12,722	9%	28%
11	APN	14,845	20%	48%
12	HEN	5,835	13%	34%



**Metric Backbone**

Ta		USN	HCN	HBFN	C-Elegans	BKF	ARP-IPP	ARP-PIP
net	NET	0.6175	0.7165	1	0.1691	0.7474	0.7374	0.6265
col	MB	0.2335	0.1318	0.4783	0.1261	0.2482	0.2832	0.2866
net	UMB	0.2335	0.0	0.0	0.1115	0.1412	0.1359	0.0913
net								

ARP-Keywords	WordNet	SCN	APN	HEN
0.9925	0.4247	0.6517	0.6696	0.5062
0.1950	0.1269	0.6165	0.6077	0.4600
0.0093	0.0214	0.6165	0.6077	0.4600

work; APN - LANL co-citation  
citation scientific collaboration



# automatic fact checking: can machines determine truth?

from data in Wikipedia

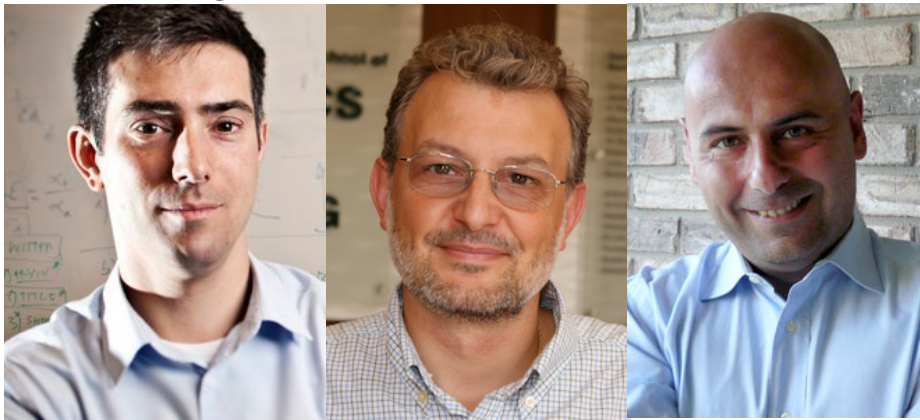


4.0 million “things” with  
470 million “facts”.



Giovanni Luca  
Ciampaglia

Using *metric closure*, predicts democrat  
from republican politician with very high  
accuracy (also decent for geography,  
capitals director, academy awards)



Johan Bollen

Fil Menczer

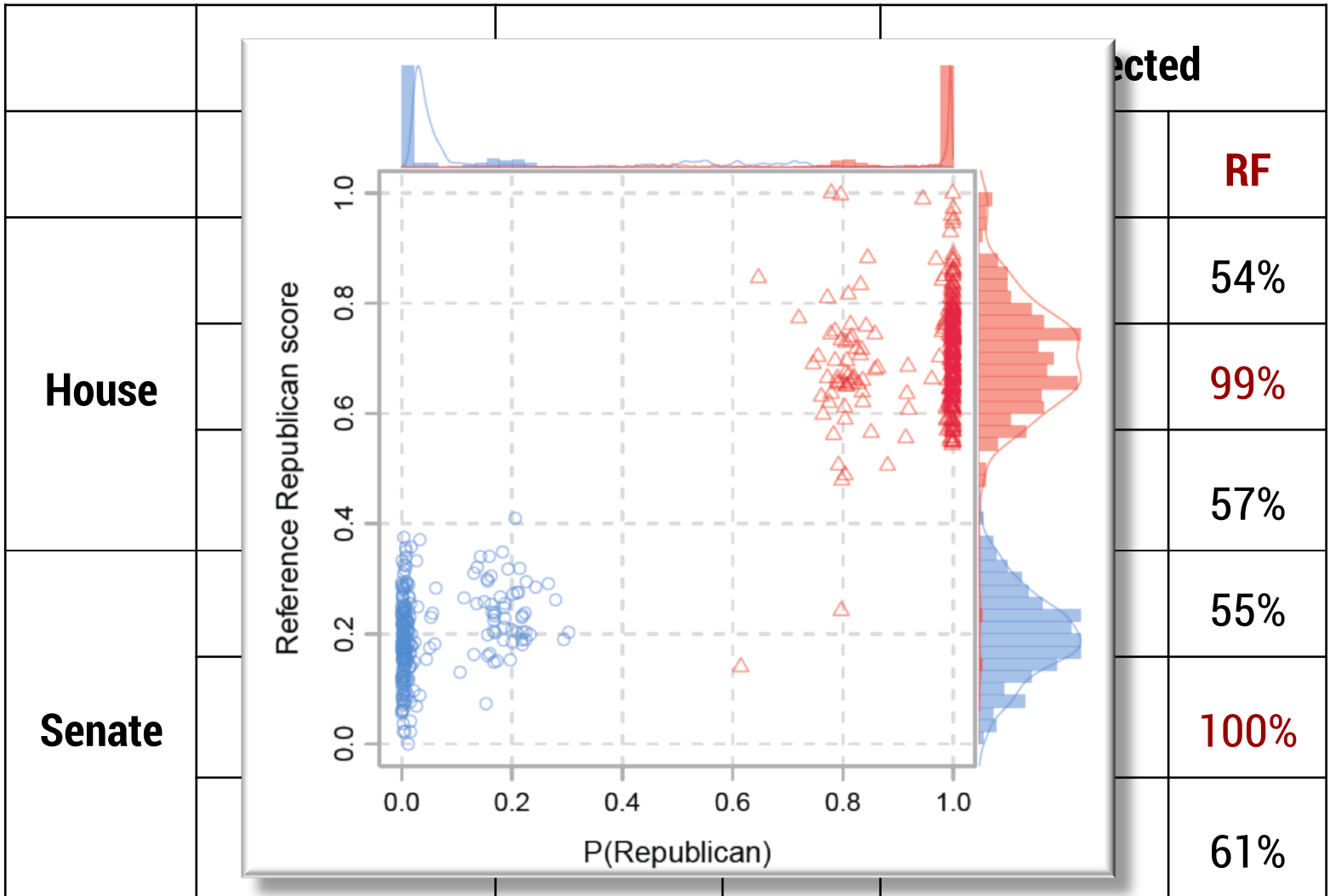
Alessandro  
Flammini

Massive graph ( $V$   
 $= 3.14M$ ;  $E = 23M$ )  
is extremely  
semimetric (98%)

[rocha@indiana.edu](mailto:rocha@indiana.edu)

<http://informatics.indiana.edu/rocha>

AUC



**Baseline:** original graph  
with no closure



## CASCI Team

and computational intelligence @ indiana university,  
<http://casci.informatics.indiana.edu>



**luis m. rocha (PI)**



- computation as a **general principle** for generating open-ended complexity (in embodied systems)
  - Turing/Von Neumann principle of universal computation/ self-replication is most fundamental principle of life: evolution via genetic variation and selection.
- **general questions** deriving from this evolutionary principle
  - How do cells and collectives of cells compute?
  - How can artificial systems evolve?
  - How does external language extend cognition?
  - Can we understand and control collective intelligence via external language? Does it evolve?
- **projects focus on language and information on networks**
  - dynamics in complex networks, automata dynamics, network topology, text mining for translational biomedicine, computational biology, agent-based modeling, artificial life, evolutionary systems, collective intelligence



and computational biology  
@ instituto gulbenkian de ciência, portugal

“Let the whole outside world consist of a long paper tape”. —John von Neumann, 1948

[rocha@indiana.edu](mailto:rocha@indiana.edu)  
<http://informatics.indiana.edu/rocha>

**Informatics**  
luis rocha 2014



selected projects – text, literature, and social media mining

## Content Extraction and Event Detection:

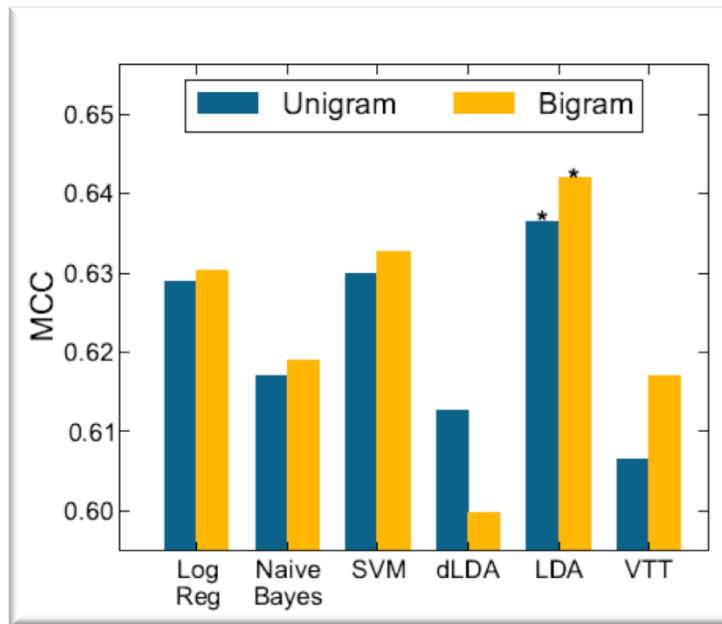
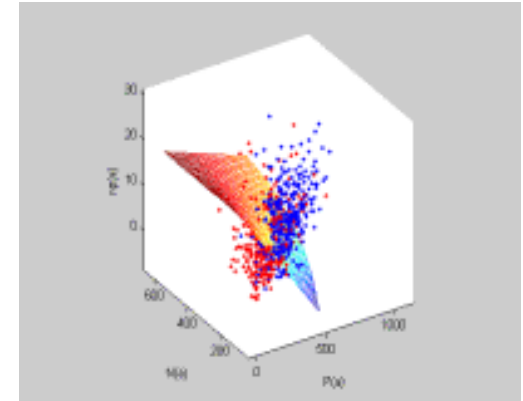
Protein-Protein Interaction, Drug-Drug-interaction, Drug Pharmacokinetics, Numerical Parameter, Event Prediction from Twitter data, Protein Structure Prediction, etc. Among top performing teams in BioCreative PPI tasks

Wu et al [2013]. *BMC Bioinformatics*, 14:35.

Wang, et al [2009]. *J. Biomedical Informatics*. 42 (4): 726-735.

Abi Haidar, A et al. [2008] *Genome Biology* 9(Suppl 2):S11.

Verspoor, K., et al [2005]. *BMC Bioinformatics*, 6(Suppl 1):S20.



## Text Classification: Retrieval of relevant documents, tweets, profiles, etc.

Lourenco, et al [2011]. *BMC Bioinformatics*. 2(Suppl 8):S12

Abi-Haidar & L.M. Rocha [2011]. *Evol. Intel.* 4(2):69-80.

Kolchinsky, et al [2010]. *Trans. Comp. Bio. Bioinf.*, 7(3):400-411

Kolchinsky et al [2013]. *PSB* 18:409-420.

Kolchinsky et al [2014]. *PLoS One*. In Press.

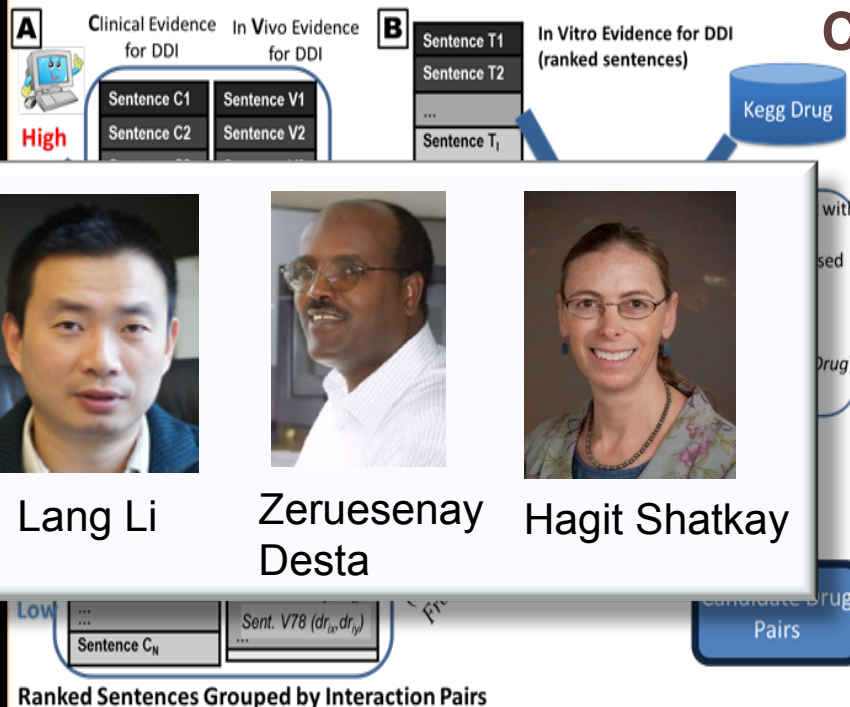
classification and extraction of evidence from the Bibliome

**threat to public health:** drug-drug interaction (DDI) is major cause of *adverse drug reaction* (ADR).  $\approx$  195K/year hospitalizations and 74K/year emergency room visits in US. Expected to increase with *polypharmacy*.

**DDI experimental evidence:** *in vitro*, molecular interactions within cell; *in vivo*, whether molecular interaction impacts drug exposure in humans; *clinical*, whether DDI changes human response to drugs (efficacy, ADR).



**Knowledge gaps:** missing evidence of any of the three types. **Need to link evidence** of molecular to clinical DDI, at collective intelligence level



Lang Li



Zeruesenay  
Desta



Hagit Shatkay

**Corpora development and Text Classification of evidence:** abstracts and sentences with evidence of DDI of one type to fill *knowledge gaps*

Kolchinsky et al [2013]. *PSB* **18**:409-420.

Wu et al [2013]. *BMC Bioinformatics*, **14**:35. DOI:10.1186/1471-2105-14-35.

Kolchinsky et al [2014]. *PLOS One*. In Press.

Wu et al [2014]. *J. Biomed. Informatics*, Submitted.

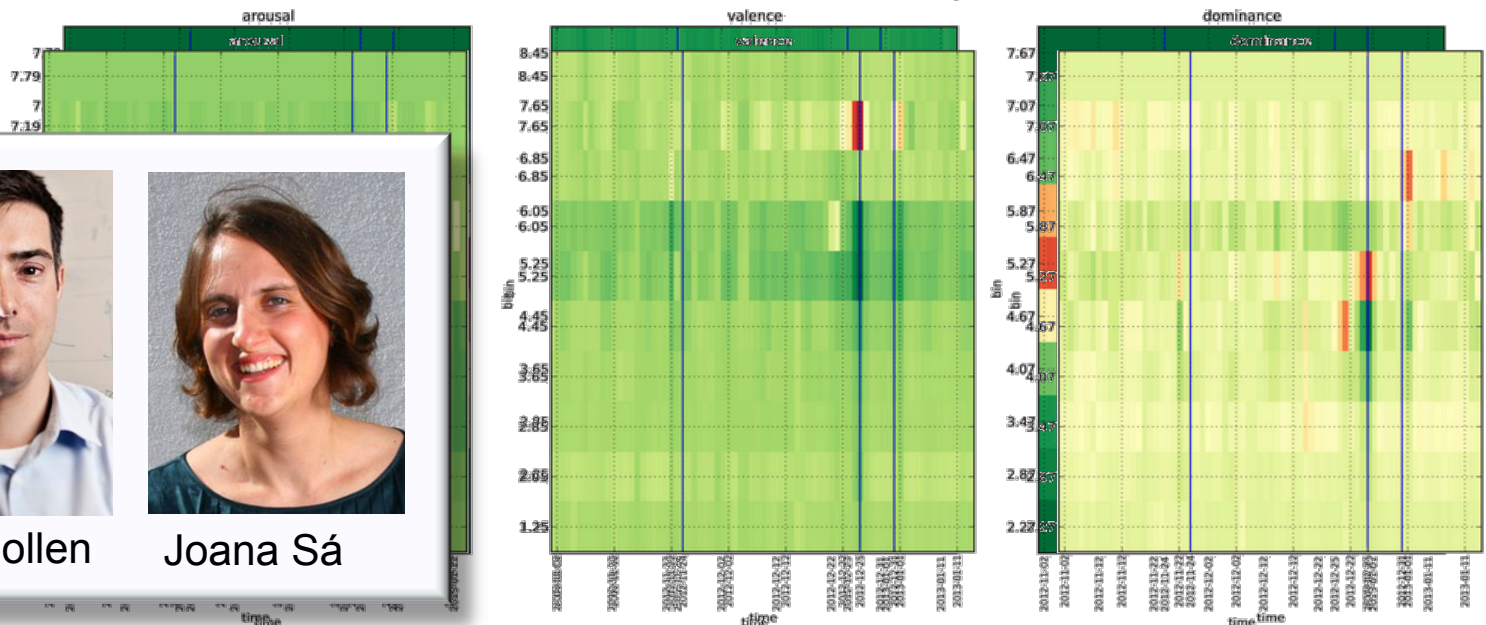
# studying collective behavior

## eigenmoods from text records on social media

- text as external component of collective behavior
  - patterns of text used by collectives
    - rather than individual brains and genes
  - e.g. Collective mood behavior on Twitter
    - Sex search data
    - Situations with divergent collective moods

### ANEW Twitter measurements 2012

#### Reconstructed matrix after first singular vector removed



Johan Bollen

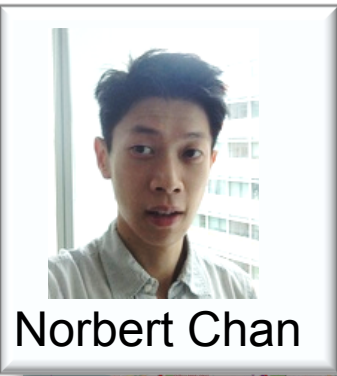


Joana Sá

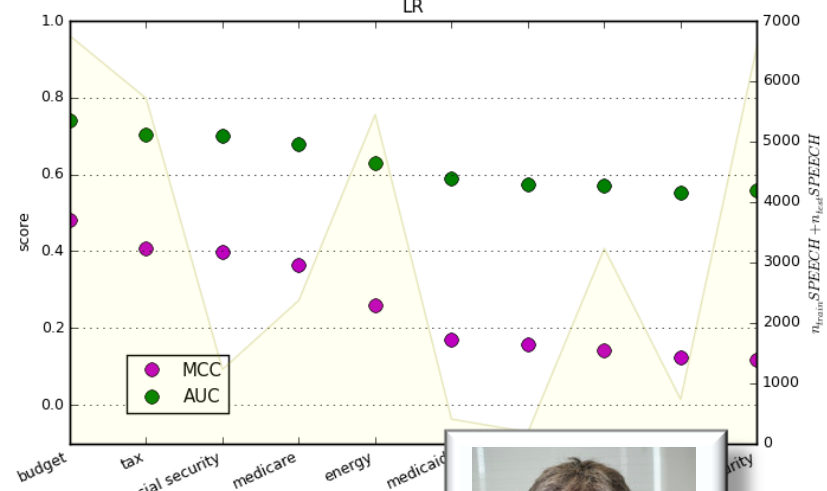
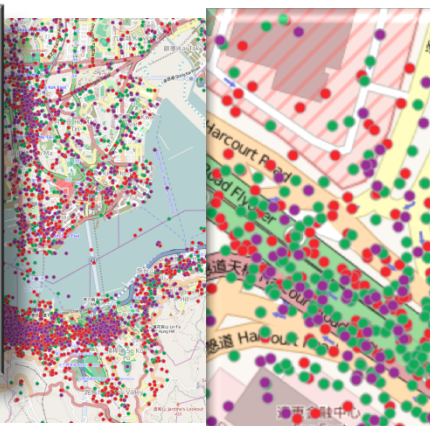


Selected projects

Political activity: social unrest via Instagram and polarization in congress



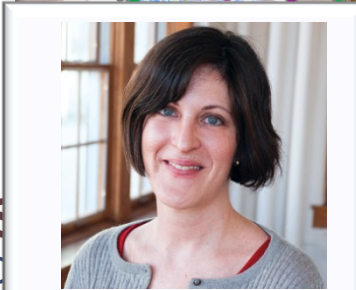
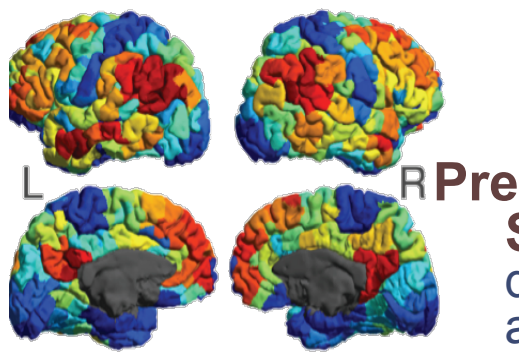
Norbert Chan



Informatics  
luis rocha 2014

Time-series Classification, RNA E

Huang, C-F, et al [2007]. *Evolution*  
Abi-Haidar & Rocha [2011]. *Evol.*  
Francisco, Wood, Sabanovic & Rocha



Selma Sabanovic



Olaf Sporns



works using statistical prediction  
theory

Kolchinsky et al [2014]. *Frontiers*

Kolchinsky & Rocha [2011]. *ECAL 2011*: 423-430.

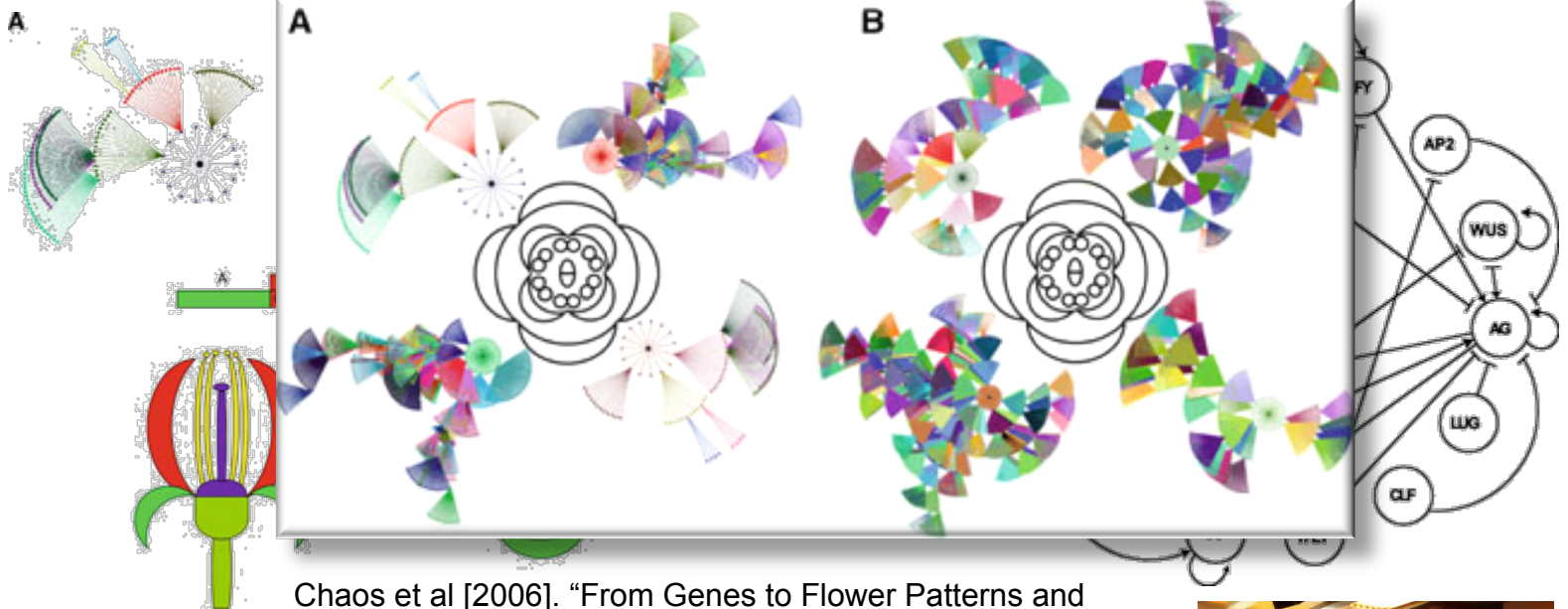


# self-organization and the cybernetics of life

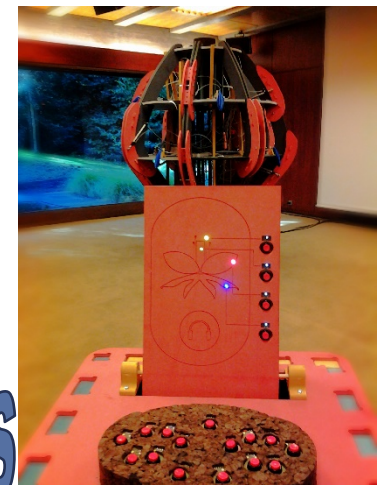
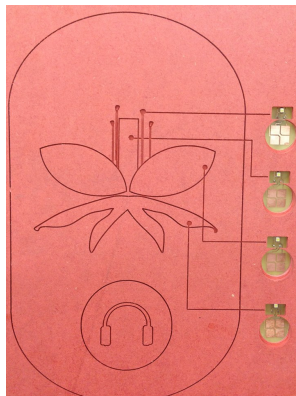
## Boolean networks, sound, art, and education



Informatics  
luis rocha 2014



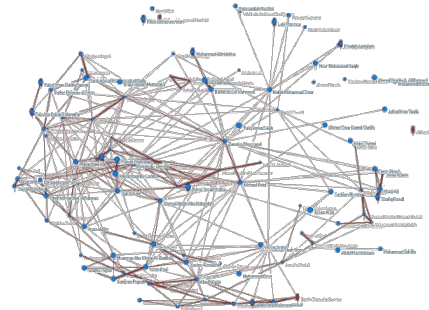
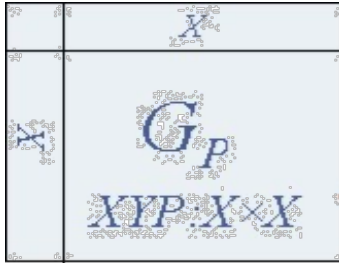
Chaos et al [2006]. "From Genes to Flower Patterns and Evolution: Dynamic Models of Gene Regulatory Networks". *Journal of Plant Growth Regulation*. 25(4): 278-289



# Musical Morphogenesis

metric closure – APSP (Dijkstra)

Proximity

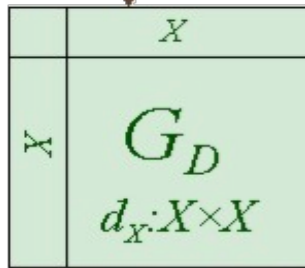


Rocha, Luis M. [2002]. In: *Soft Computing Agents*. V. Loia (Ed.):137-163.

Rocha, L.M. et al [2005]. *IEEE Web Intelligence (WI'05)*: 565-571.

shortest-path where *path length* is sum of edges

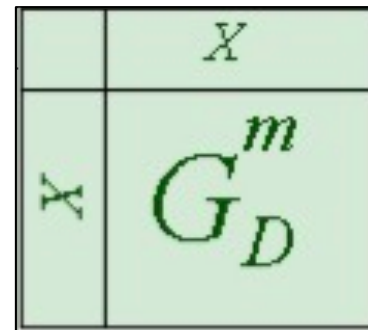
$$d_X(x_i, x_j) = \frac{1}{XYP(x_i, x_j)} - 1$$



Distance

Metric Closure (min/+)

= metric  
> semi-metric



(Shortest Path)

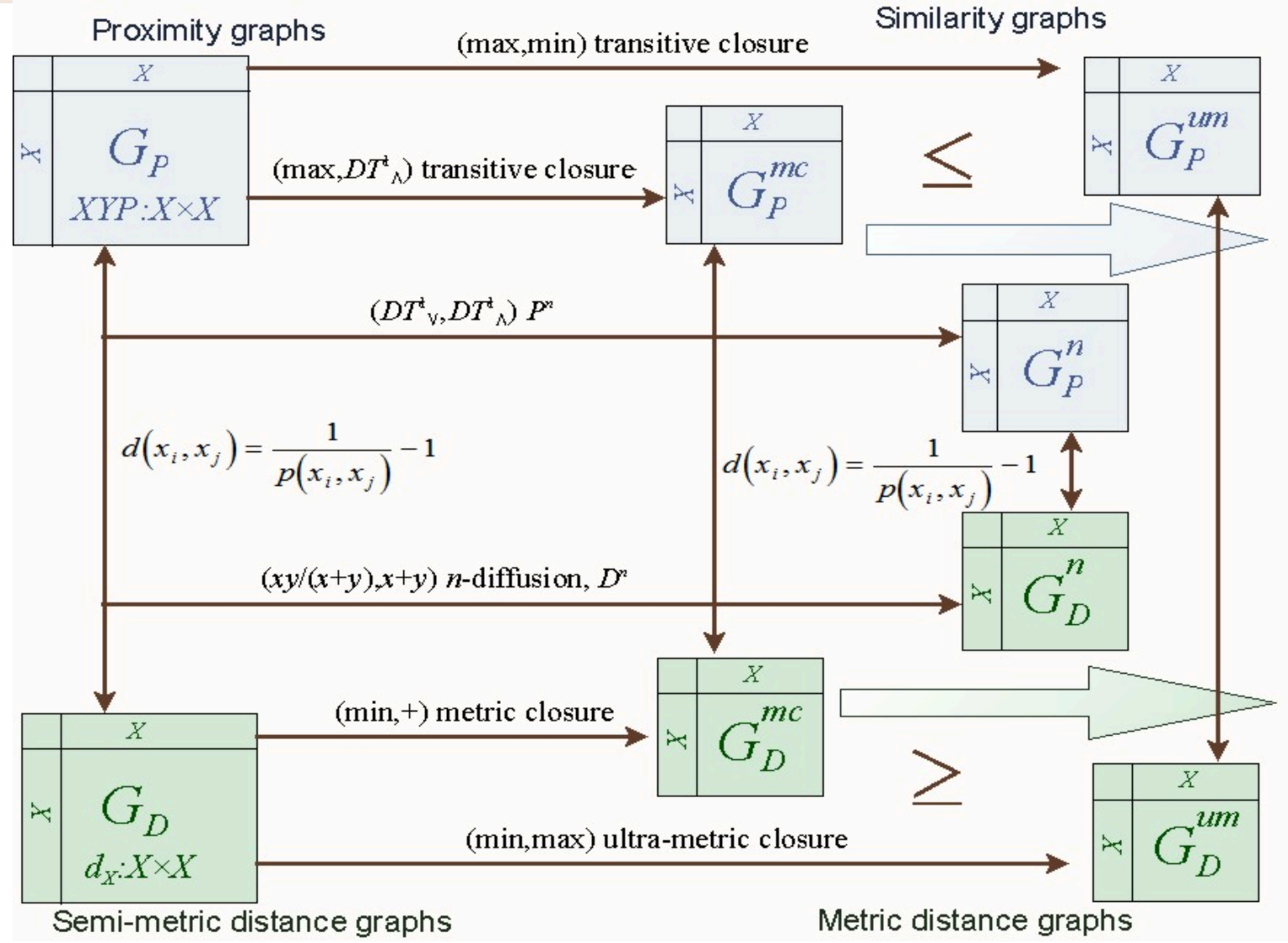
$$b \downarrow i, j = d \downarrow i, j$$

$$s \downarrow i, j = d \downarrow i, j$$

Measures of semi-metric behavior



all cases



Informatics  
luis rocha 2014



intuition about latent associations



close to metric



more semi-metric







intuition



ultra-metric closure



weak weakest link



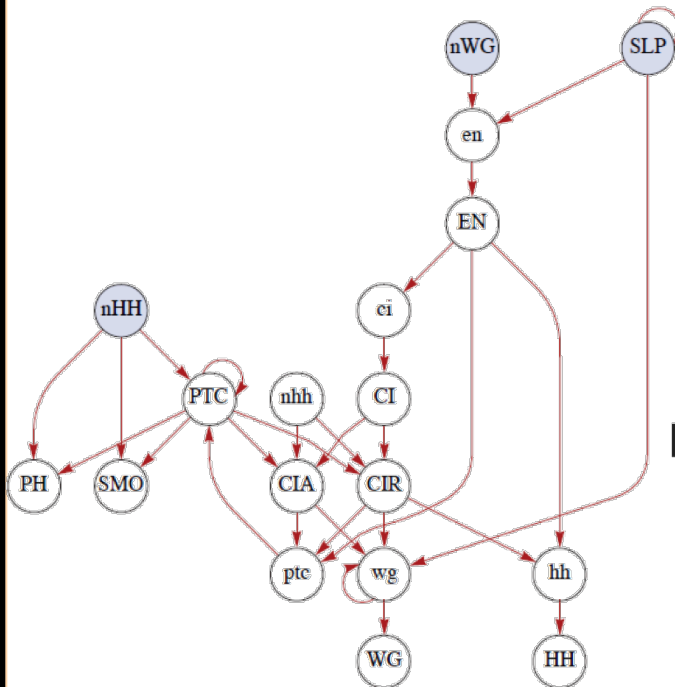
strong weakest link



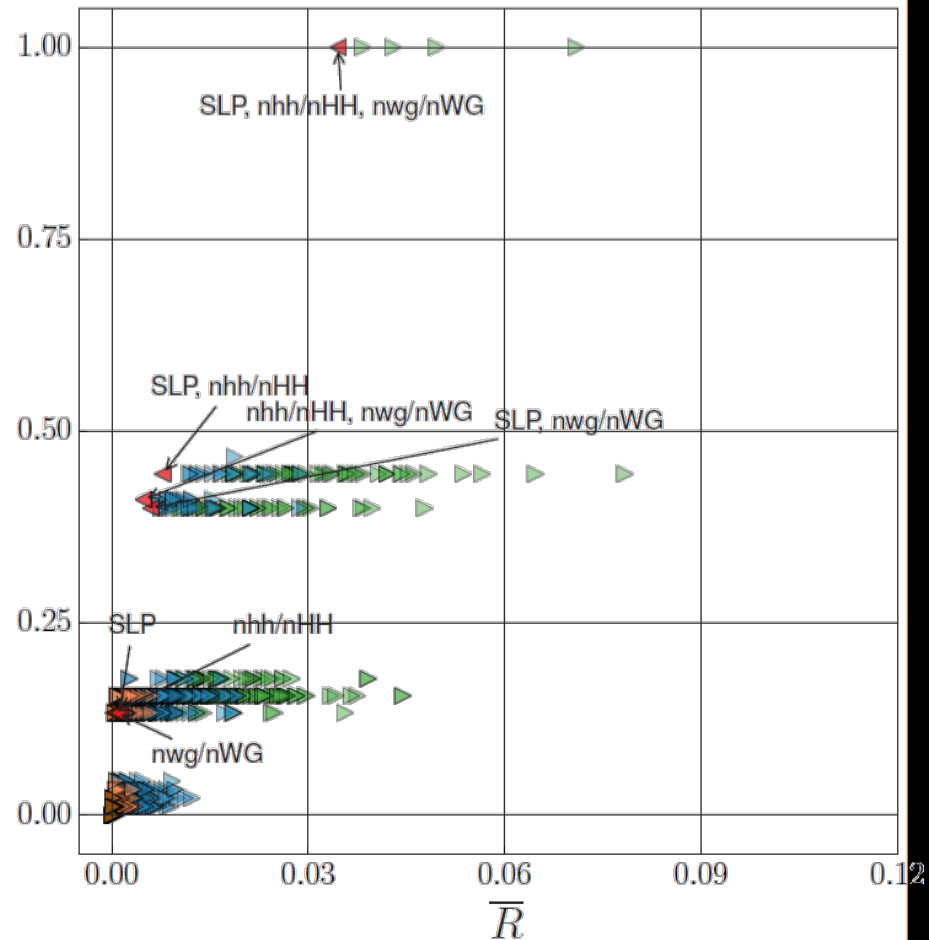
Tends to link items very strongly  
Greater distortion of original graph  
Worse for inference

# Can structural controllability uncover control?

## Drosophila model (Albert et al)



$\bar{A}$



4 nodes predicted by structural control:

- $\{SLP, nWG, nhh/nHH, PH\}$ ,
- $\{SLP, nWG, nhh/nHH, SMO\}$ ,
- $\{SLP, nWG, nhh/nHH, CIR\}$ ,
- $\{SLP, nWG, nhh/nHH, CIA\}$

Gates & Rocha [2014]. *ALIFE 2014*: 429-430.

Gates & Rocha [2014]. *In preparation*..

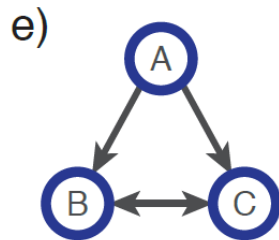
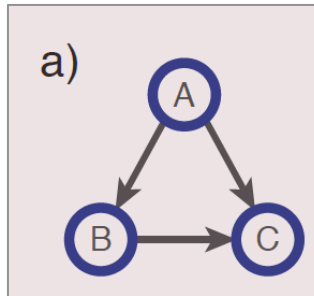
[rocha@indiana.edu](mailto:rocha@indiana.edu)

<http://informatics.indiana.edu/rocha>

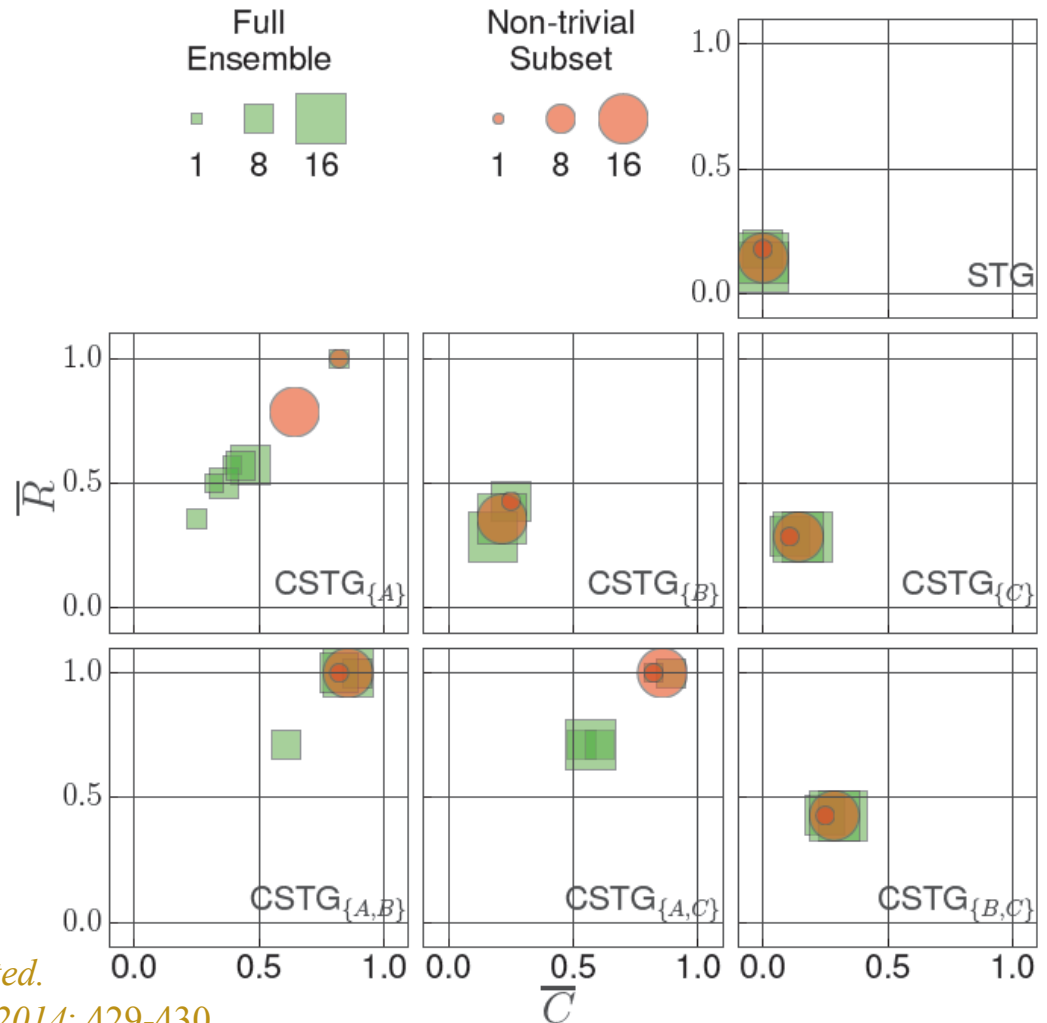
# Can structural controlability uncover control?

In presence of dynamics

Consider small network motifs:



The larger the **canalization**, the less predictable is structural control



Gates & Rocha [2014]. *Submitted*.  
Gates & Rocha [2014]. *ALIFE 2014*: 429-430.