Bioinformatics 3

V7 – Function Annotation, Gene Regulation

Mon., Nov 5, 2012

Network Meta-Growth

Q: When I find a new protein and its (already known) partners
in an experiment and I add that to a database,
do I get a scale-free network?

Which proteins are in databases??? <=> the experimentally accessibly ones!!!

→ costs for the experiment
→ experience required for purification, methods, analysis…
→ existing assays for similar proteins
→ personal interests
(to get funding: preference for {cancer, HIV, Alzheimer…})

Higher probability to find proteins related to known ones
<=> growing network with preferential attachment

What Does a Protein Do?

BRENDA

The Comprehensive Enzyme Information System

TU Braunschweig Dept. of Bioinformatics

EC **Explorer** [SEARCH] [BROWSE]

- **1 Oxidoreductases** *(4042 organisms)*
- **2 Transferases** *(3198 organisms)*
  - 2.1 Transferring one-carbon groups *(615 organisms)*
    - 2.1.1 Methyltransferases *(514 organisms)*
    - 2.1.2 Hydroxymethyl-, formyl- and related transferases *(82 organisms)*
    - 2.1.3 Carboxy- and carbamoyltransferases *(105 organisms)*
    - 2.1.4 Amidinotransferases *(32 organisms)*
      - 2.1.4.1 glycine amidinotransferase *(17 organisms)*
      - 2.1.4.2 *scyllo*-inosamine-4-phosphate amidinotransferase *(15 organisms)*
  - 2.2 Transferring aldehyde or ketonic groups *(91 organisms)*
  - 2.3 Acyltransferases *(930 organisms)*
  - 2.4 Glycosyltransferases *(925 organisms)*
  - 2.5 Transferring alkyl or aryl groups, other than methyl groups *(547 organisms)*
  - 2.6 Transferring nitrogenous groups *(377 organisms)*
  - 2.7 Transferring phosphorus-containing groups *(1343 organisms)*
  - 2.8 Transferring sulfur-containing groups *(276 organisms)*
  - 2.9 Transferring selenium-containing groups *(6 organisms)*
- **3 Hydrolases** *(4453 organisms)*
- **4 Lyases** *(2145 organisms)*
- **5 Isomerases** *(849 organisms)*
- **6 Ligases** *(686 organisms)*

Enzyme **C**lassification scheme (from http://www.brenda-enzymes.org/)

MIPS FunCat

Classification Browser from http://mips.gsf.de/projects/funcat

Digging Deeper

- MIPS Functional Catalogue
  - 01 METABOLISM
    - 01.01 amino acid metabolism
    - 01.02 nitrogen, sulfur and selenium metabolism
    - 01.03 nucleotide/nucleoside/nucleobase metabolism
    - 01.04 phosphate metabolism
    - 01.05 C-compound and carbohydrate metabolism
    - 01.06 lipid, fatty acid and isoprenoid metabolism
      - 01.06.02 membrane lipid metabolism
        - 01.06.02.01 phospholipid metabolism
        - 01.06.02.02 glycolipid metabolism
        - 01.06.02.03 sphingolipid metabolism
      - 01.06.05 fatty acid metabolism
      - 01.06.06 isoprenoid metabolism
      - 01.06.10 regulation of lipid, fatty acid and isoprenoid metabolism
    - 01.07 metabolism of vitamins, cofactors, and prosthetic groups
    - 01.08 metabolism of intercellular mediators
    - 01.20 secondary metabolism
    - 01.25 extracellular metabolism
  - 02 ENERGY
  - 04 STORAGE PROTEIN
  - 10 CELL CYCLE AND DNA PROCESSING
  - 11 TRANSCRIPTION
  - 12 PROTEIN SYNTHESIS
  - 14 PROTEIN FATE (folding, modification, destination)
  - 16 PROTEIN WITH BINDING FUNCTION OR COFACTOR REQUIREMENT (structural or catalytic)
  - 18 REGULATION OF METABOLISM AND PROTEIN FUNCTION
  - 20 CELLULAR TRANSPORT, TRANSPORT FACILITIES AND TRANSPORT ROUTES
  - 30 CELLULAR COMMUNICATION/SIGNAL TRANSDUCTION MECHANISM

Details and Proteins

01 METABOLISM [Eukaryota Bacteria Archaea]
01.06 lipid, fatty acid and isoprenoid metabolism [Eukaryota Bacteria Archaea]
01.06.02 membrane lipid metabolism [Eukaryota Bacteria Archaea]
01.06.02.01 phospholipid metabolism [Eukaryota Bacteria Archaea]

DETAILED RESULTS:

Number:	01.06.02.01
Description:	phospholipid metabolism
Explanation	-
EC number:	-
Taxonomy:	[Eukaryota Bacteria Archaea]
Reference Link:	-
Reference PMID:	-
Funlink:	-
GO mapping:	GO:0008654

Manually Annotated Proteins:

Organism	Proteins
Helicobacter pylori KE26695	HP0190; HP0700; HP0215; HP0737; HP1071; HP0961; HP0871; HP1016; HP0201; ***Co-annotated-FunCats***
Saccharomyces cerevisiae	
Neurospora crassa	
Listeria monocytogenes EGD	gi_16411389; gi_16410893; gi_16411991; gi_16409732; gi_16410732; gi_16410825; gi_16409367; gi_16411263; ***Co-annotated-FunCats***

http://mips.gsf.de/projects/funcat

Un-Classified Proteins?

BIOINFORMATICS

Vol. 21 Suppl. 1 2005, pages i302–i310
doi:10.1093/bioinformatics/bti1054

Whole-proteome prediction of protein function via graph-theoretic analysis of interaction maps

Elena Nabieva[1,2], Kam Jim[2], Amit Agarwal[1], Bernard Chazelle[1] and Mona Singh[1,2,]*

[1]Computer Science Department and [2]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

Received on January 15, 2005; accepted on March 27, 2005

Many **unclassified proteins**:

→ estimate: ~1/3 of the yeast proteome not annotated functionally

→ BioGRID: 4495 proteins in the largest cluster of the yeast physical interaction map.

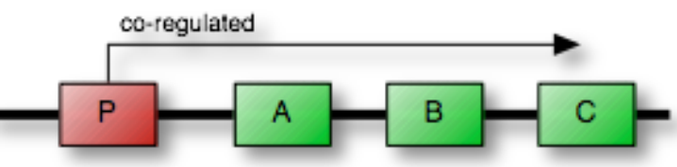
2946 have a MIPS functional annotation

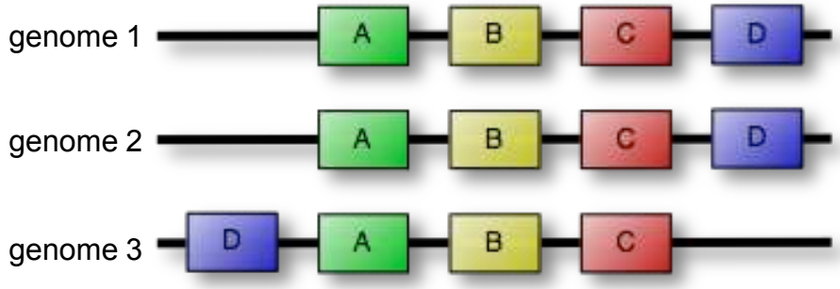
Partition the Graph

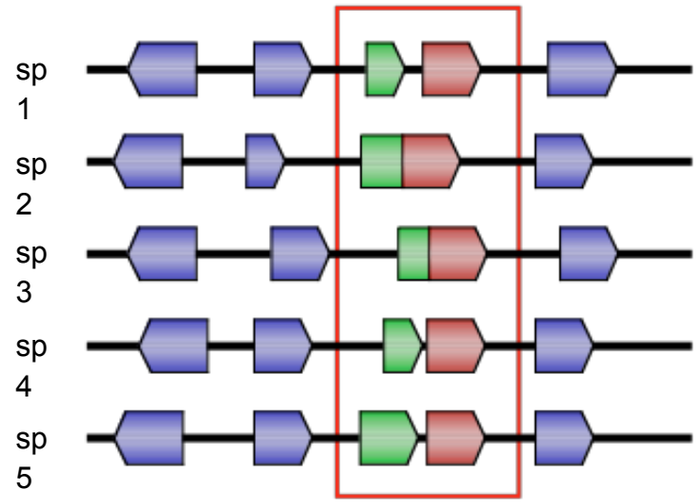
Large **PPI networks** were built from:

- HT experiments (Y2H, TAP, synthetic lethality, coexpression, coregulation, …)
- predictions (gene profiling, gene neighborhood, phylogenetic profiles, …)

→ proteins that are functionally linked







Identify **unknown functions** from **clustering** of these networks by, e.g.:

- shared interactions (similar neighborhood → power graphs)
- membership in a community
- similarity of shortest path vectors to all other proteins (= similar path into the rest of the network)

Protein Interactions

Nabieva et al used the *S. cerevisiae* dataset from GRID of 2005 (now BioGRID)
→ 4495 proteins and 12 531 physical interactions in the largest cluster

BioGRID

Search GO

Organism: Escherichia coli K12

General Repository for Interaction Datasets

home | help / support | contribute | downloads | mirrors | about us

About BioGRID

The Biological General Repository for Interaction Datasets (BioGRID) database (http://www.thebiogrid.org) was developed to house and distribute collections of protein and genetic interactions from major model organism species. BioGRID currently contains over 198 000 interactions from six different species, as derived from both high-throughput studies and conventional focused studies. Through comprehensive curation efforts, BioGRID now includes a virtually complete set of interactions reported to date in the primary literature for both the budding yeast Saccharomyces cerevisiae and the fission yeast Schizosaccharomyces pombe. A number of new features have been added to the BioGRID including an improved user interface to display interactions based on different attributes, a mirror site and a dedicated interaction management system to coordinate curation across different locations. The BioGRID provides interaction data with monthly updates to Saccharomyces Genome Database, Flybase and Entrez Gene. Source code for the BioGRID and the linked Osprey network visualization system is now freely available without restriction.

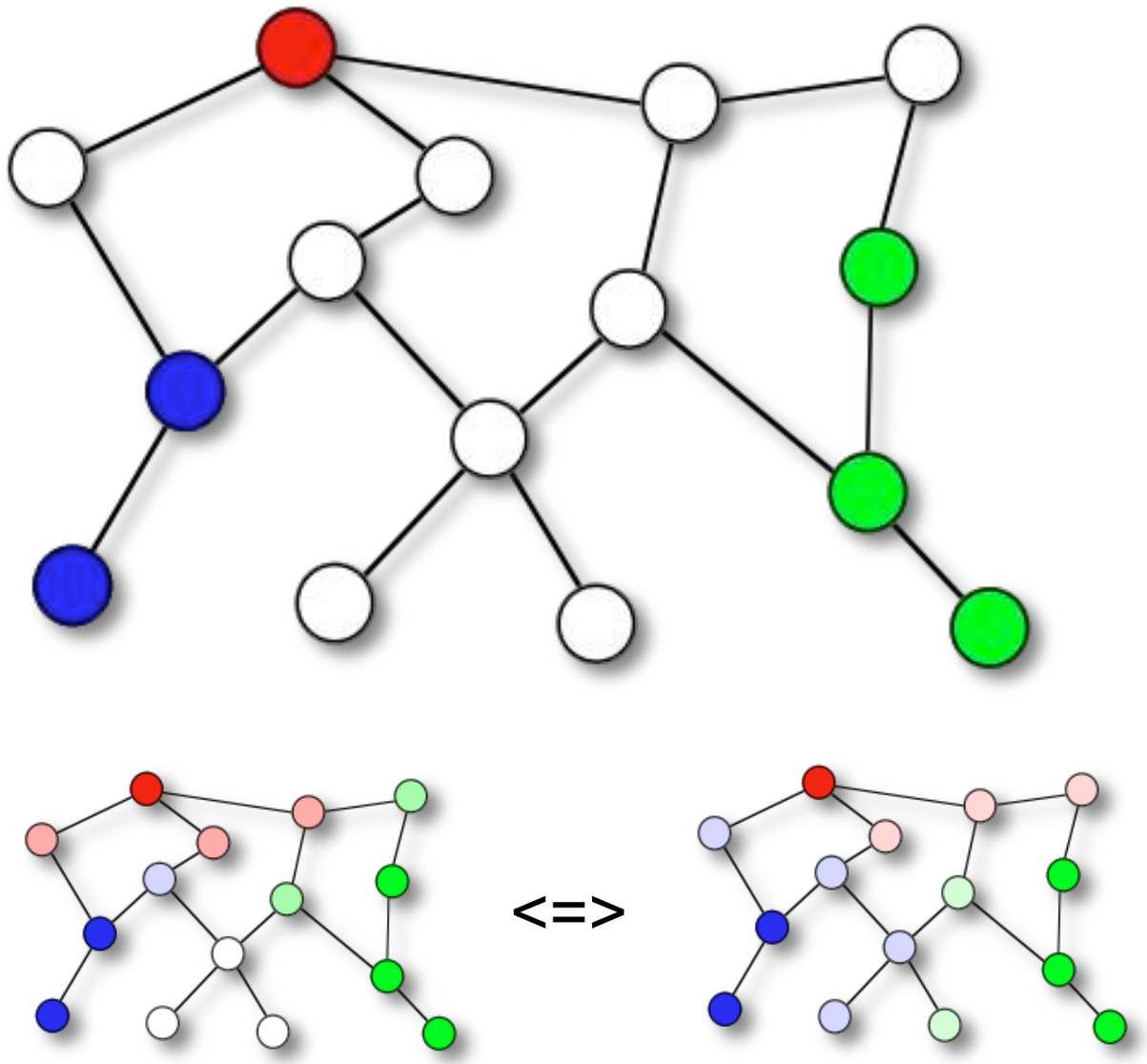
BioGRID Links

- Arabidopsis Information Resource
- BioPIXIE
- Biotechnology and Biological Sciences Research Council (BBSRC)
- Canadian Institutes of Health Research (CIHR)
- Cytoscape
- Database of Interacting Proteins
- Entrez-Gene
- Flybase
- Gene DB
- Gene Ontology
- Germ Online

http://www.thebiogrid.org/about.php

Function Annotation

Task: **predict** function (= functional annotation) for a protein from the **available** annotations



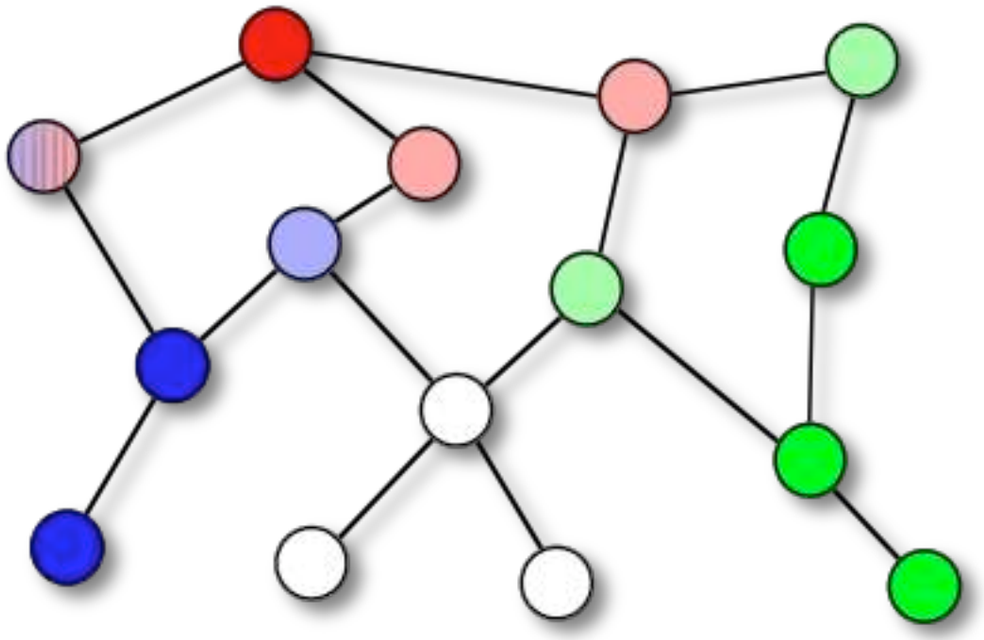
Similar:
How to **assign colors** to the white nodes?

Use information on:
- distance to colored nodes
- local connectivity
- reliability of the links
- …

Algorithm I: Majority

Schwikowski, Uetz, and Fields, " A network of protein–protein interactions in yeast" *Nat. Biotechnol.* **18** (2000) 1257

Consider all neighbors and **sum** up how often a certain **annotation occurs**
→ score for an annotation = count among the direct neighbors
→ take the 3 most frequent functions



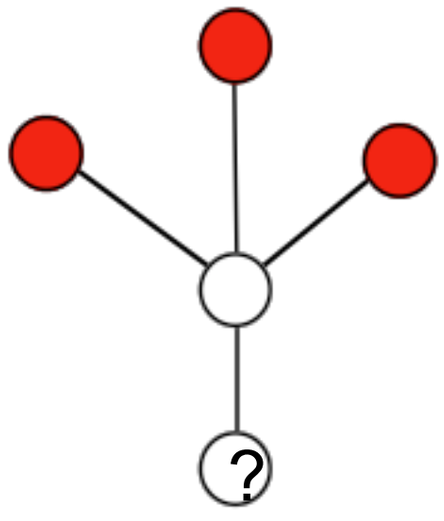
Majority makes only limited use of the local connectivity
→ cannot assign function to next-neighbors

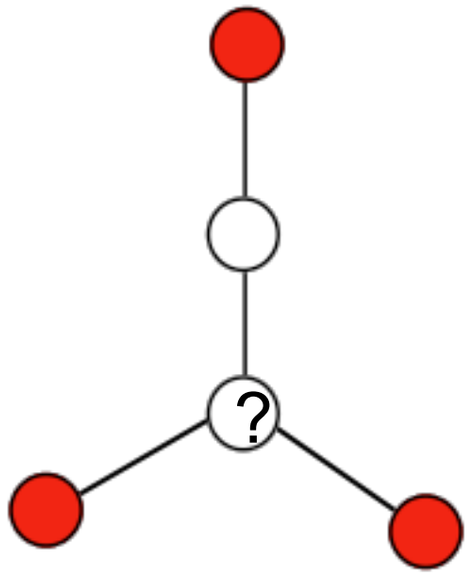
For weighted graphs:
→ weighted sum

Extended Majority: Neighborhood

Hishigaki, Nakai, Ono, Tanigami, and Takagi, "Assessment of prediction accuracy of protein function from protein–protein interaction data", *Yeast* **18** (2001) 523

Look for **overrepresented** functions within a given **radius** of 1, 2, or 3 links
→ use as function score the value of a χ^2–test





Neighborhood does not consider local network topology

Both examples are treated **identical** with $r = 2$

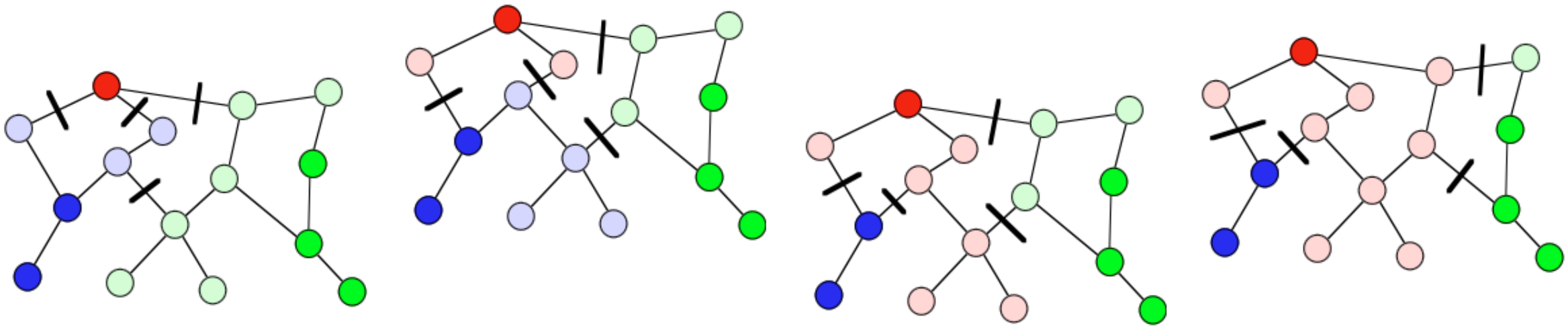
Neighborhood can not (easily) be generalized to weighted graphs!

Minimize Changes: GenMultiCut

Karaoz, Murali, Letovsky, Zheng, Ding, Cantor, and Kasif, "Whole-genome annotation by using evidence integration in functional-linkage networks" PNAS **101** (2004) 2888

"Annotate proteins so as to **minimize** the number of times that **different** functions are associated with **neighboring** proteins"

→ generalization of the multiway *k*-cut problem for weighted edges,
can be stated as an integer linear program (ILP)



Multiple possible solutions → scores from **frequency** of annotations

Nabieva *et al*: FunctionalFlow

Extend the idea of **"guilty by association"**

→ each annotated protein is a source of "function"-flow

→ simulate for a few time steps

→ choose the annotation with the highest accumulated flow

Each node *u* has a reservoir $R_t(u)$, each edge a capacity constraint (weight) $w_{u,v}$

Initially: $R_0^a(u) = \begin{cases} \infty, & \text{if } u \text{ is annotated with } a, \\ 0, & \text{otherwise.} \end{cases}$ and $g_0^a(u,v) = 0$

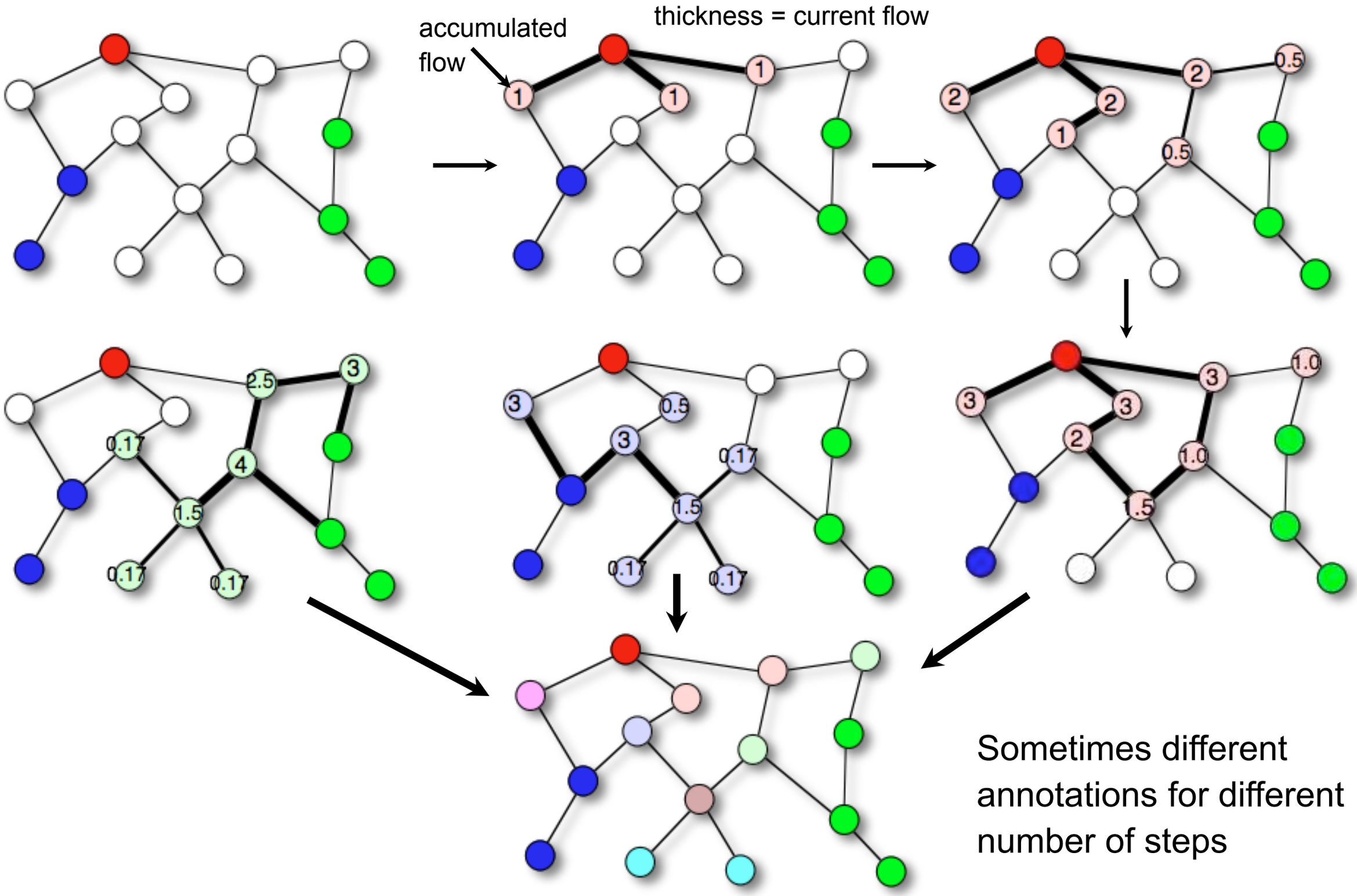
Then: **downhill flow** with capacity contraints

$$g_t^a(u,v) = \begin{cases} 0, & \text{if } R_{t-1}^a(u) < R_{t-1}^a(v) \\ \min\left(w_{u,v}, \frac{w_{u,v}}{\sum_{(u,y)\in E} w_{u,y}}\right), & \text{otherwise.} \end{cases}$$

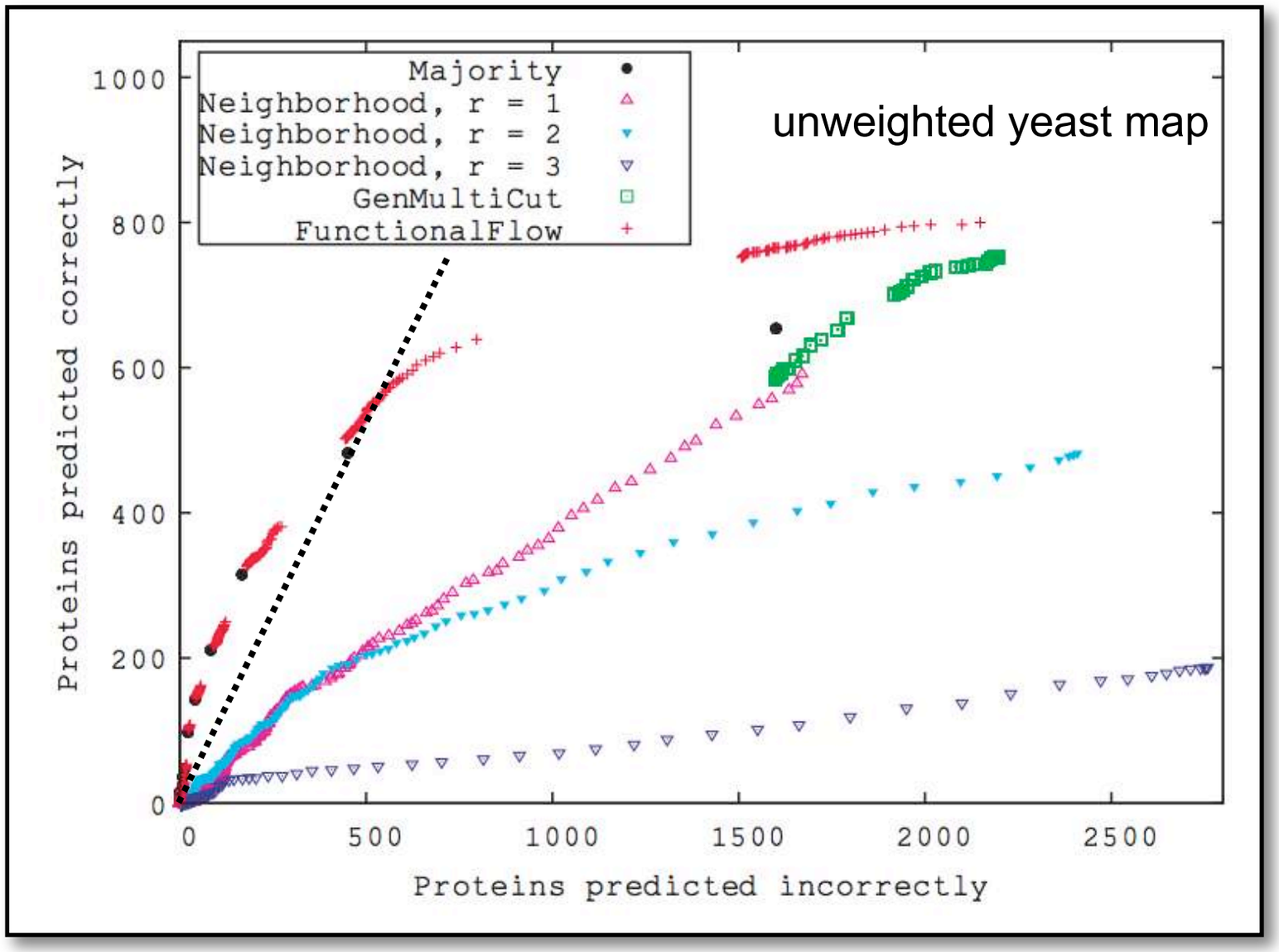
Score from accumulated in-flow: $f_a(u) = \sum_{t=1}^{d} \sum_{v:(u,v)\in E} g_t^a(v,u)$

Nabieva et al, Bioinformatics 21 (2005) i302

An Example


accumulated
flow
thickness = current flow
Sometimes different
annotations for different
number of steps

Comparison

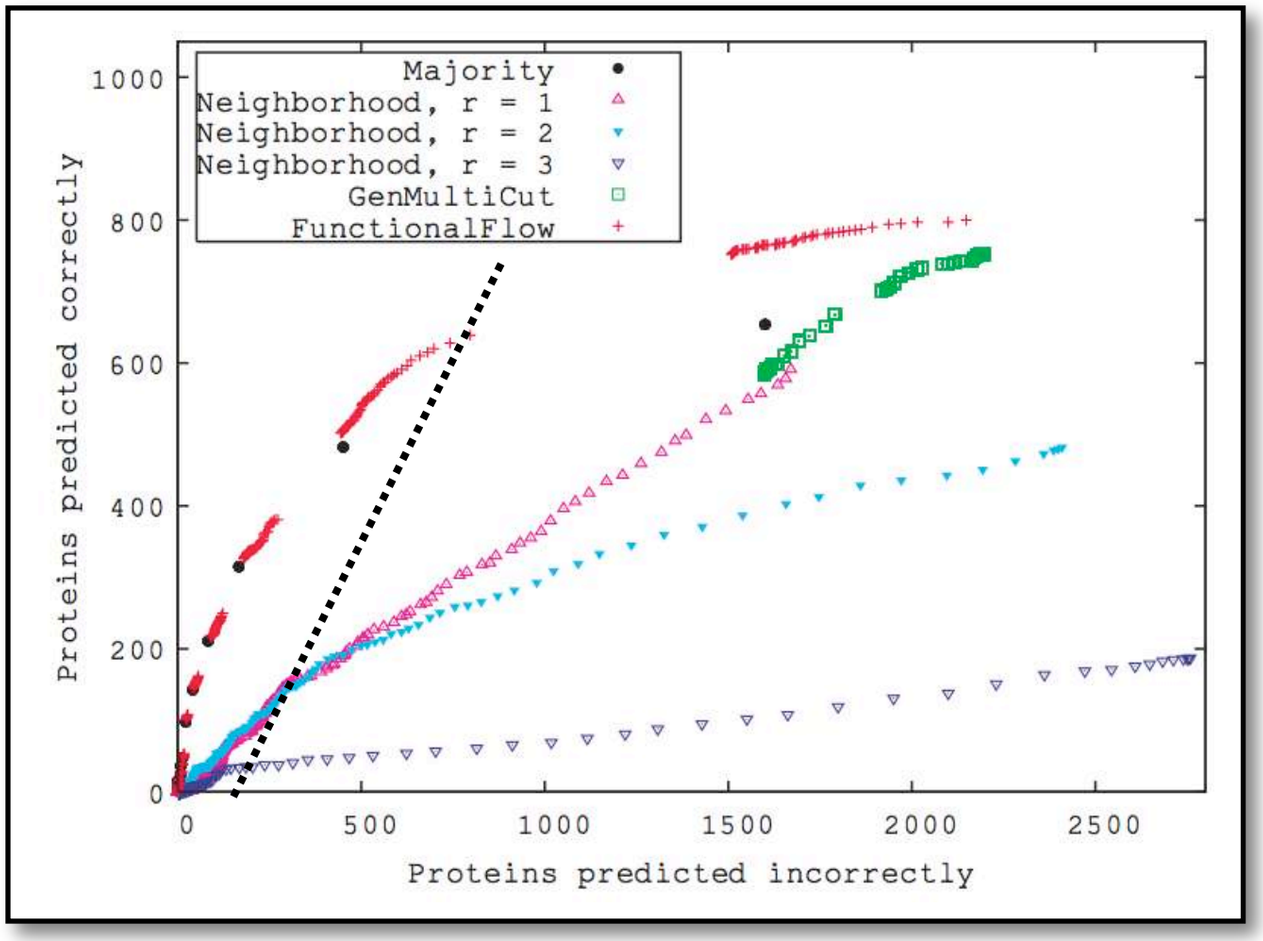


For FunctionalFlow:
six propagation steps
(diameter of the yeast
network ≈ 12)

Change **score threshold** for accepting annotations → ratio **TP/FP**
→ **FunctionalFlow** performs **best** in the high-confidence region
→ many false predictions!!!

Nabieva et al, Bioinformatics 21 (2005) i302

Comparison Details



Multiple runs (solutions) of FunctionalFlow
(with slight random perturbations of the weights)
→ increases prediction accuracy

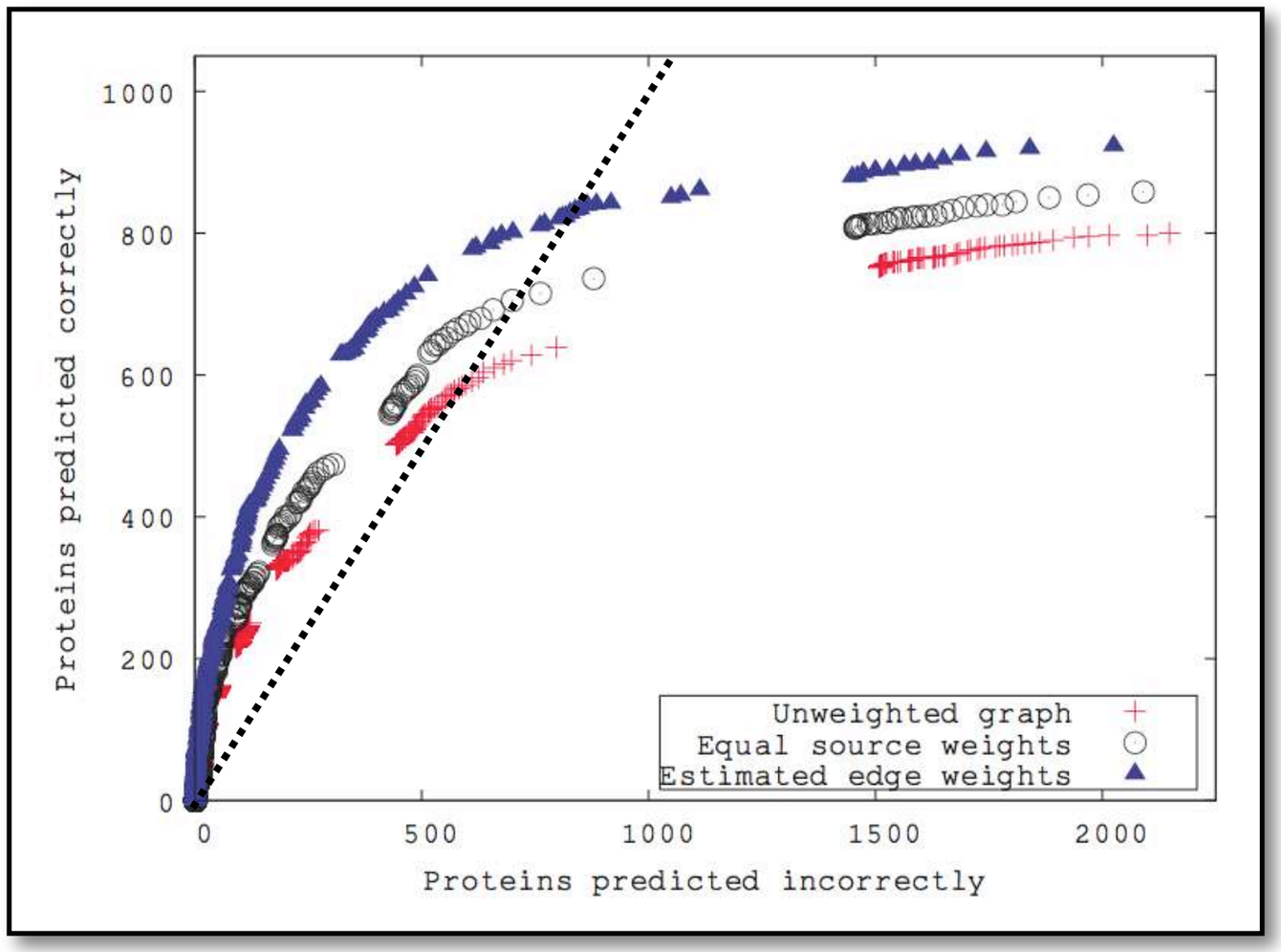
Majority vs. Neighborhood @ $r = 1$
→ counting neighboring annotations is more effective than χ^2-test

Neighborhood with $r = 1$ or 2 comparable to FunctionalFlow
for high-confidence region, performance decreases with increasing r
→ **bad** idea to **ignore** local connectivity

Nabieva et al, Bioinformatics 21 (2005) i302

Weighted Graphs

Performance of FunctionalFlow with differently weighted data:



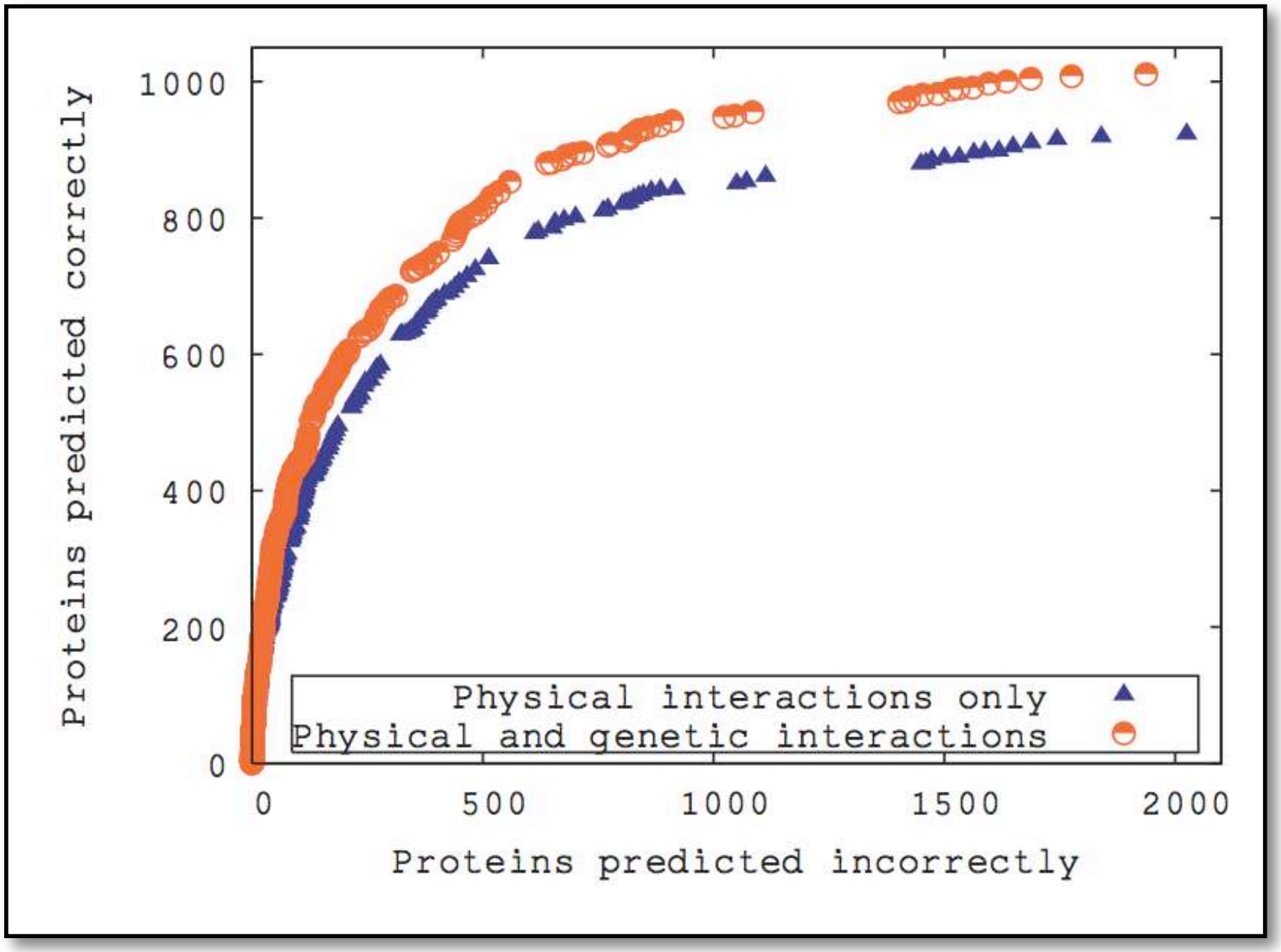
Compare:

- unweighted
- weight 0.5 per experiment
- weight for experiments according to (estimated) reliability

Largest improvement
→ individual experimental reliabilities

Nabieva et al, Bioinformatics 21 (2005) i302

Additional Information



(Note the clever choice of symbols in the plot…)

Use **genetic linkage** to modify the edge **weights**
→ better performance (also for Majority and GenMultiCut)

Nabieva et al, Bioinformatics 21 (2005) i302

Summary: **Static** PPI-Networks

"Proteins are **modular machines**" <=> How are they related to each other?

1) **Understand** "Networks"
   prototypes (ER, SF, …) and their properties (*P(k), C(k),* clustering, …)

2) Get the **data**
   experimental and theoretical techniques (Y2H, TAP, co-regulation, …),
   quality control and data integration (Bayes)

3) **Analyze** the data
   compare *P(k), C(k),* clusters, … to prototypes → highly modular, clustered
   with sparse sampling → PPI networks are not scale-free

4) **Predict** missing information
   network structure combined from multiple sources → functional annotation

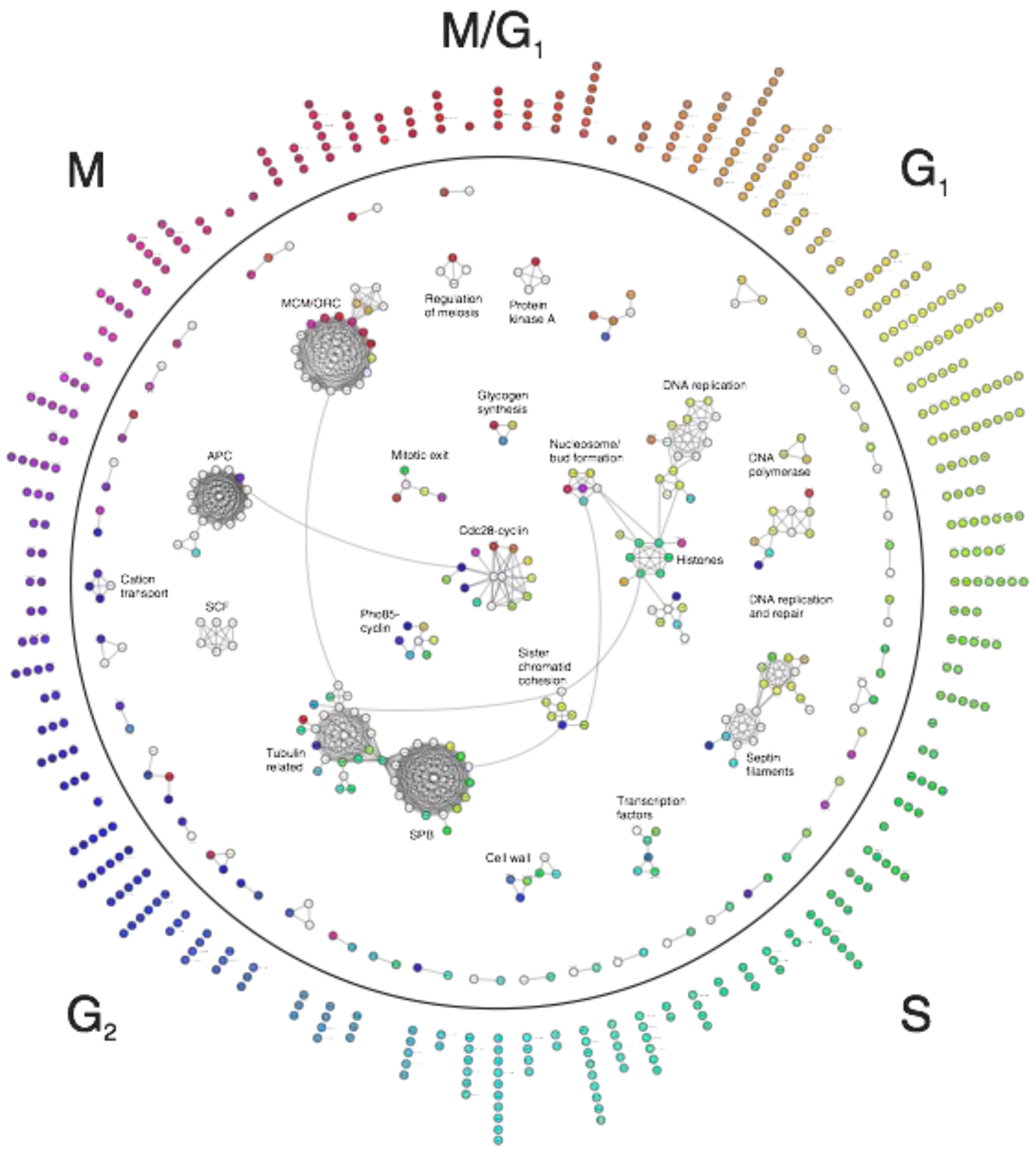
Next step: environmental changes, cell cycle
→ **changes** (dynamics) in the PPI network – **how and why?**

Turn, Turn, Turn…

From Lichtenberg et al, Science 307 (2005) 724:

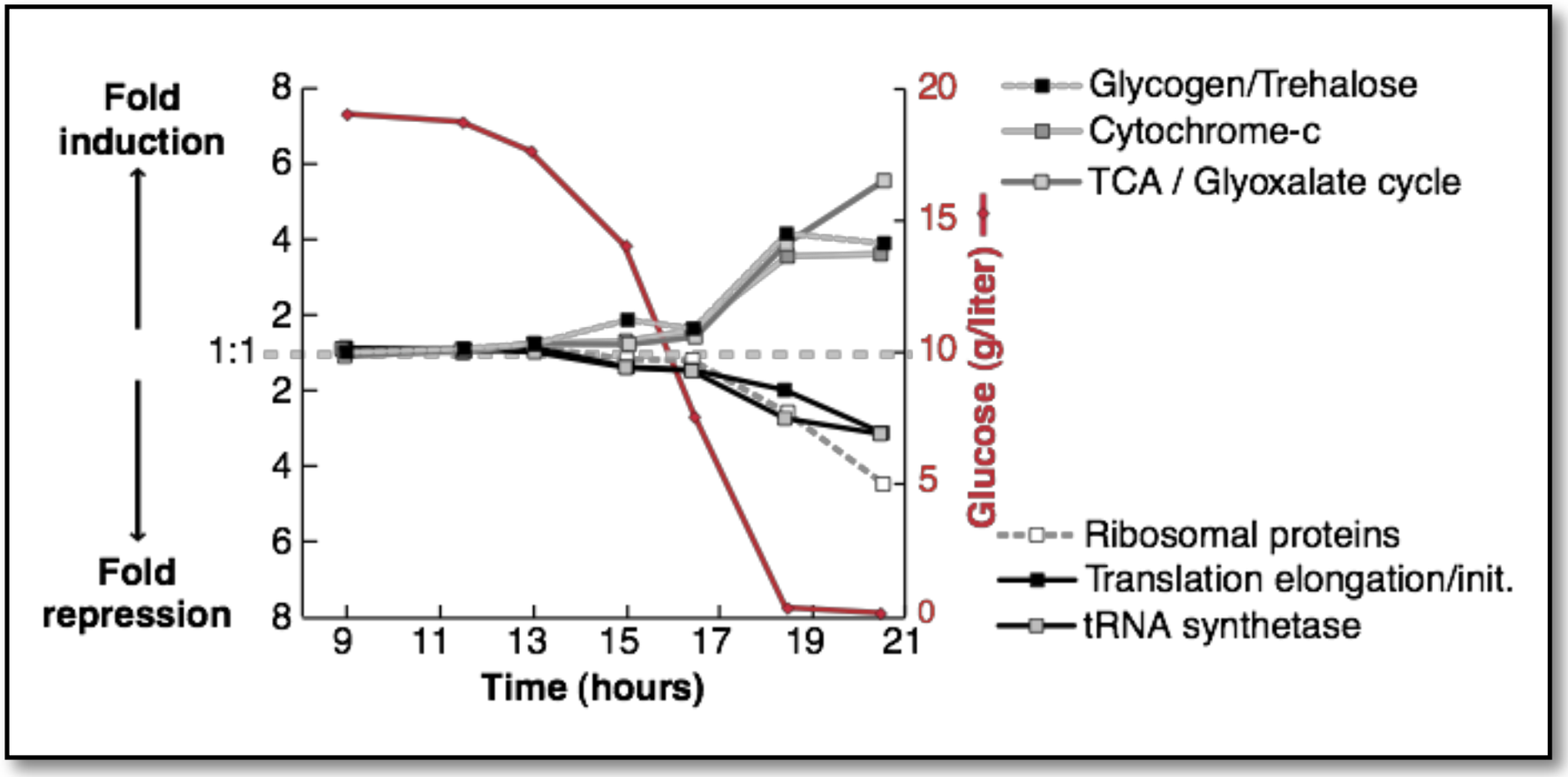
→ certain proteins only occur during well-defined phases in the cell cycle

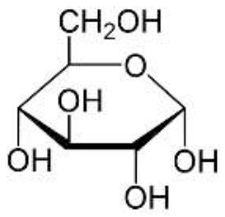
→ how is protein expression regulated?

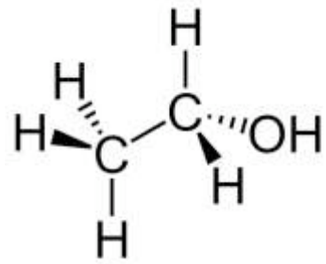


External Triggers

Re-routing of metabolic fluxes during the diauxic shift in *S. cerevisiae*
→ changes in protein abundances (measured via mRNA levels)







anaerobic fermentation:
fast growth on glucose → ethanol

Diauxic shift →

aerobic respiration:
ethanol as carbons source

Note: "quorum sensing" — different bacteria have different strategies

DeRisi et al., *Science* **278** (1997) 680

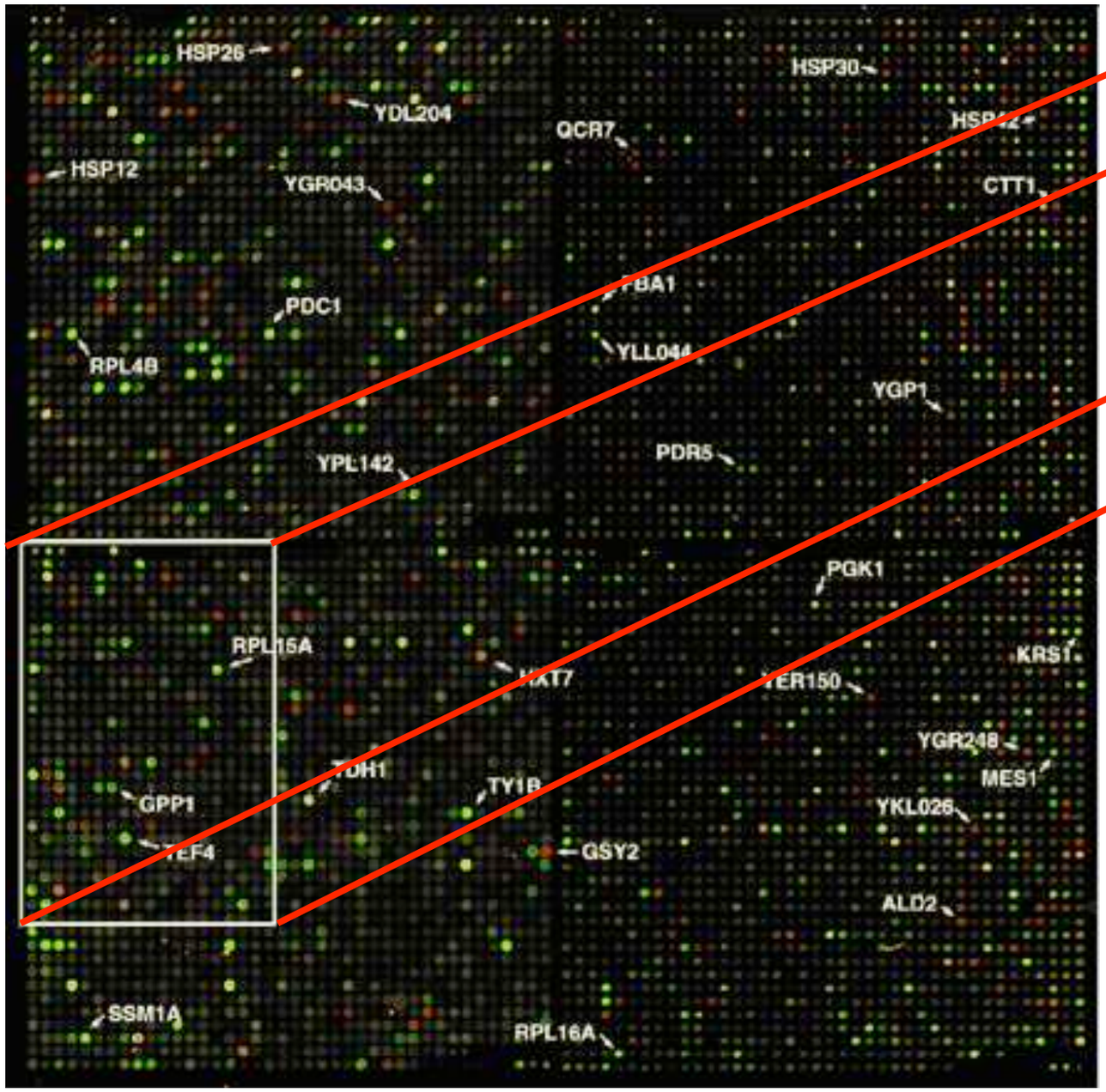
Diauxic shift affects hundreds of genes

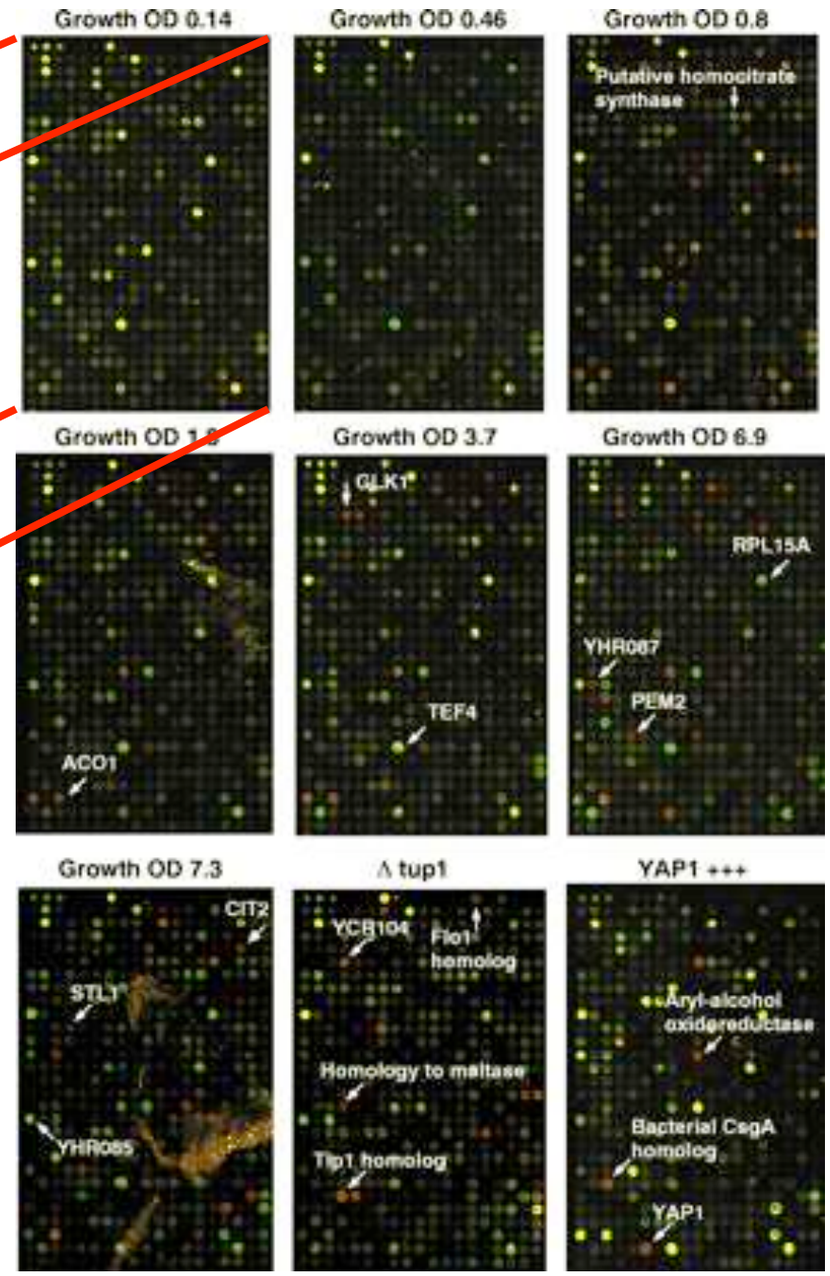
Cy3/Cy5 labels, comparison of 2 probes at 9.5 hours distance; w and w/o glucose

Optical density (OD) illustrates cell growth;

Red: genes induced by diauxic shift (710 genes 2-fold)

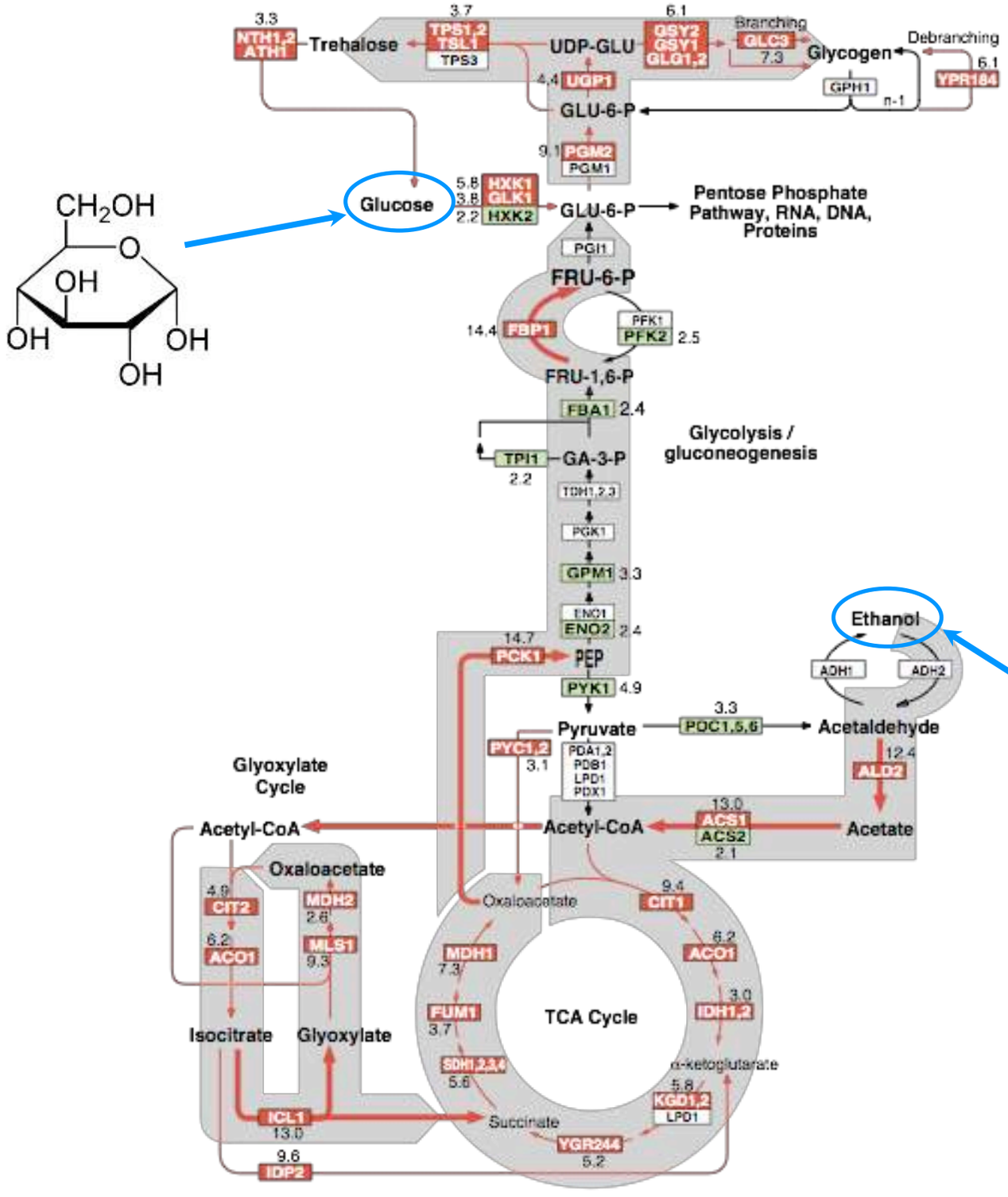
Green: genes repressed by diauxic shift (1030 genes 2-fold

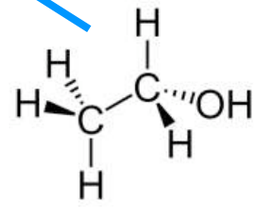




DeRisi et al., *Science* **278** (1997) 680

Flux Re-Routing





during diauxic shift:

fold change

expression increases

expression unchanged

expression diminishes

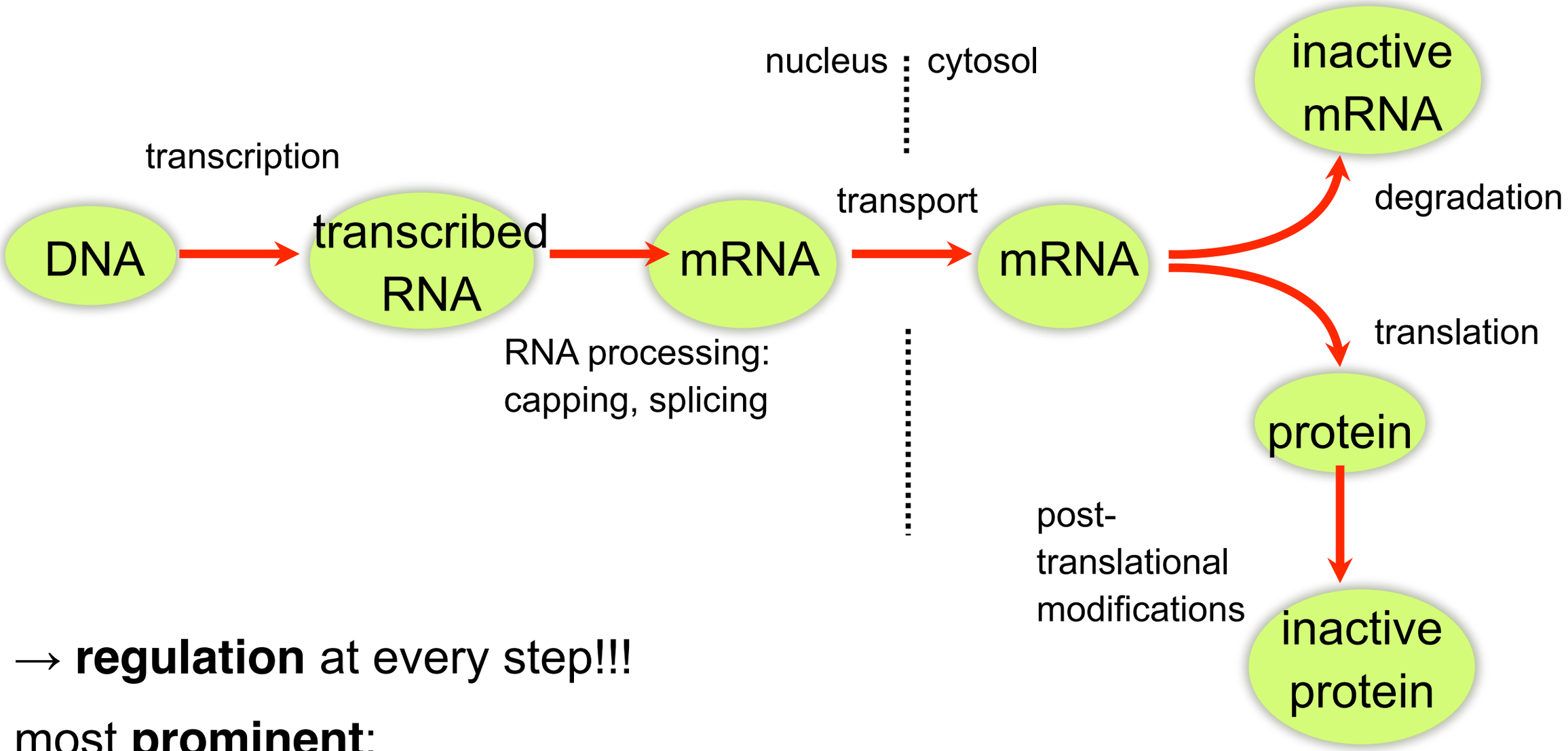
metabolic flux increases

→ **how** are these changes **coordinated?**

DeRisi et al., *Science* **278** (1997) 680

Gene Expression

Sequence of processes: from DNA to functional proteins



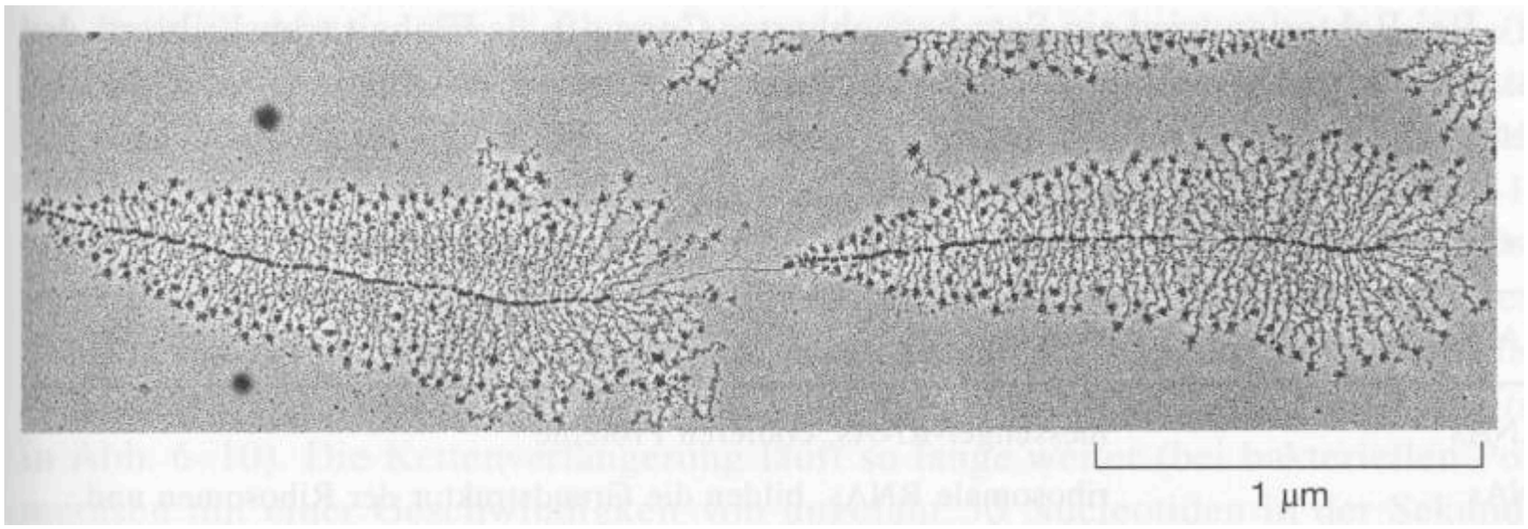
→ **regulation** at every step!!!

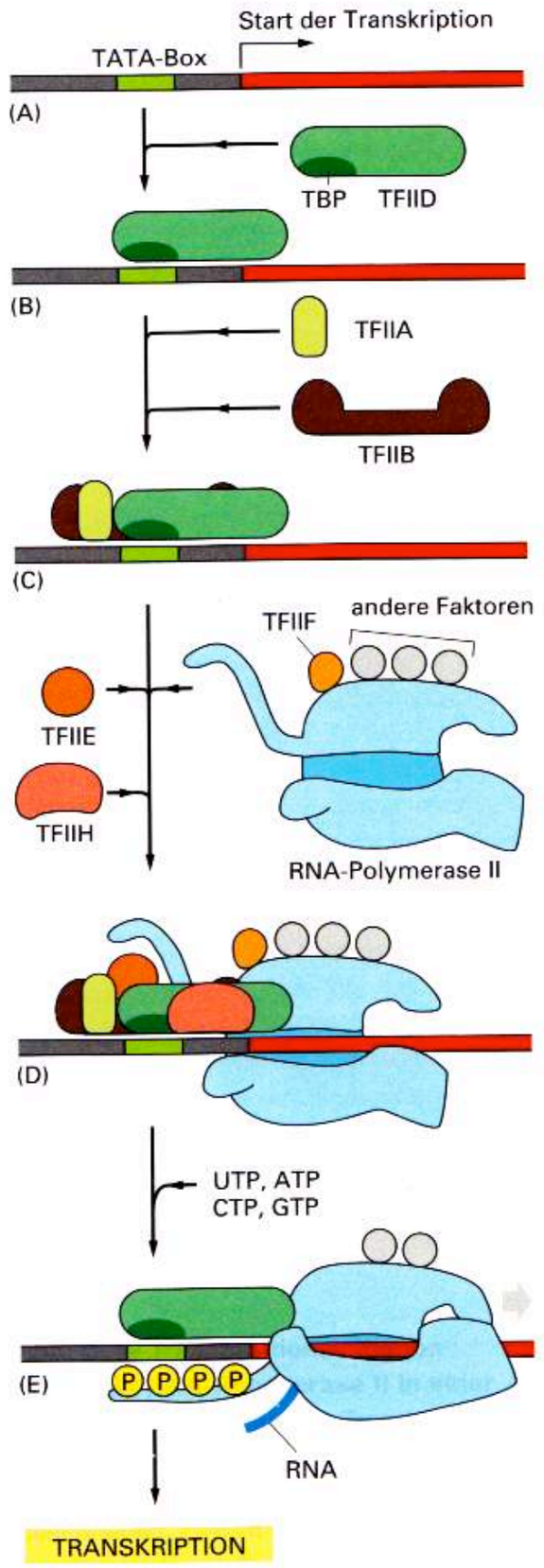
most **prominent**:
activation or repression of the transcription initiation

Transcription Initiation

In eukaryotes:

- several **general** transcription factors **have** to bind
- **specific** enhancers or repressors **may** bind
- then the RNA polymerase binds
- and starts transcription





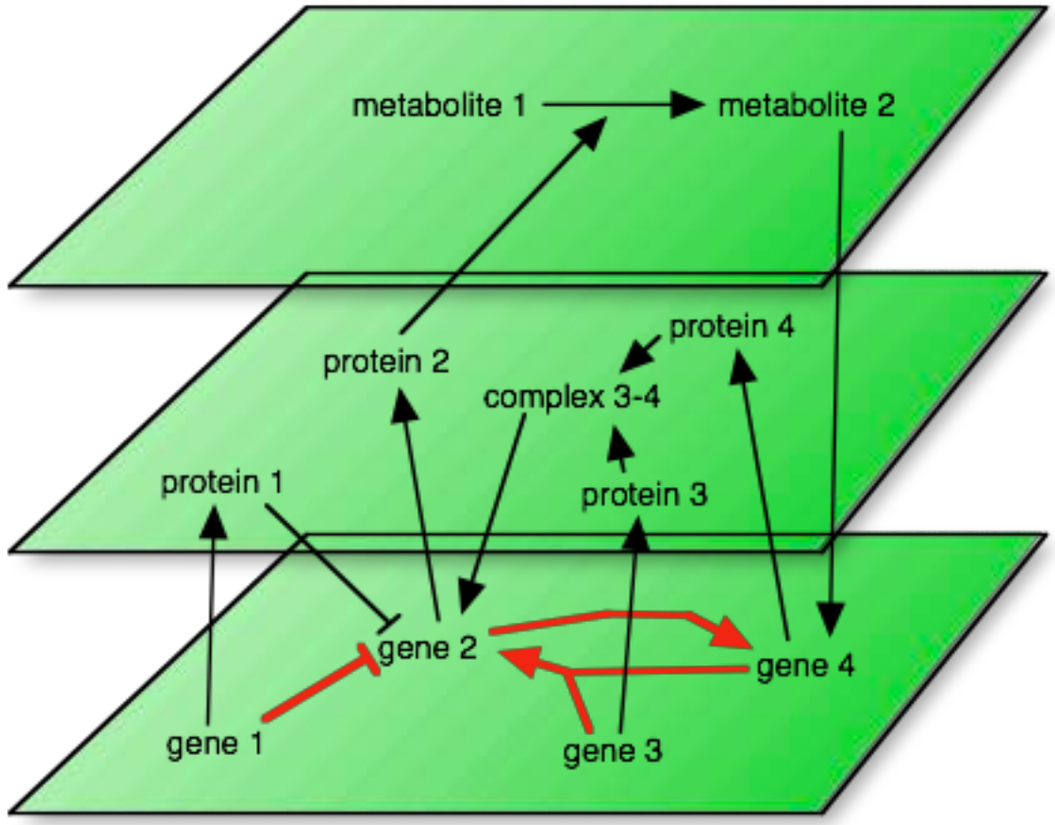
Alberts et al.
"Molekularbiologie der Zelle", 4. Aufl.

Layers upon Layers

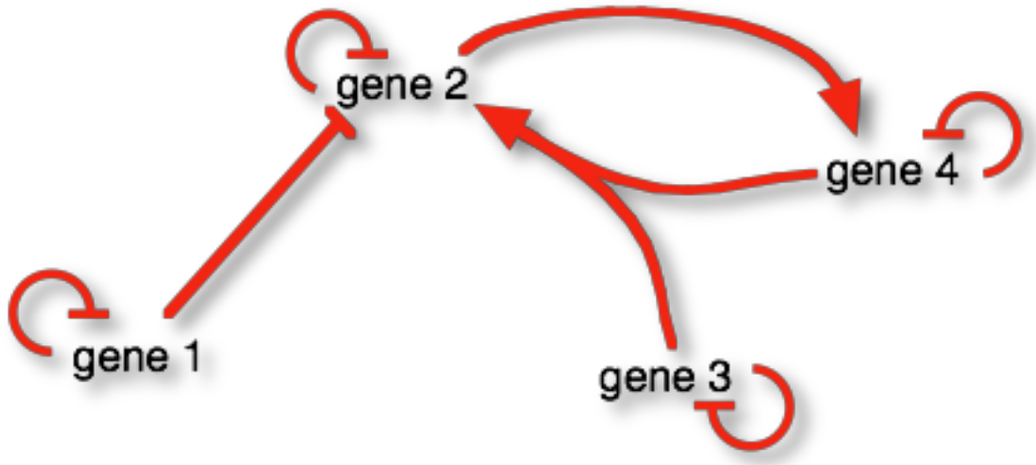
Biological regulation
via proteins and metabolites

<=>

Projected regulatory network



<=>



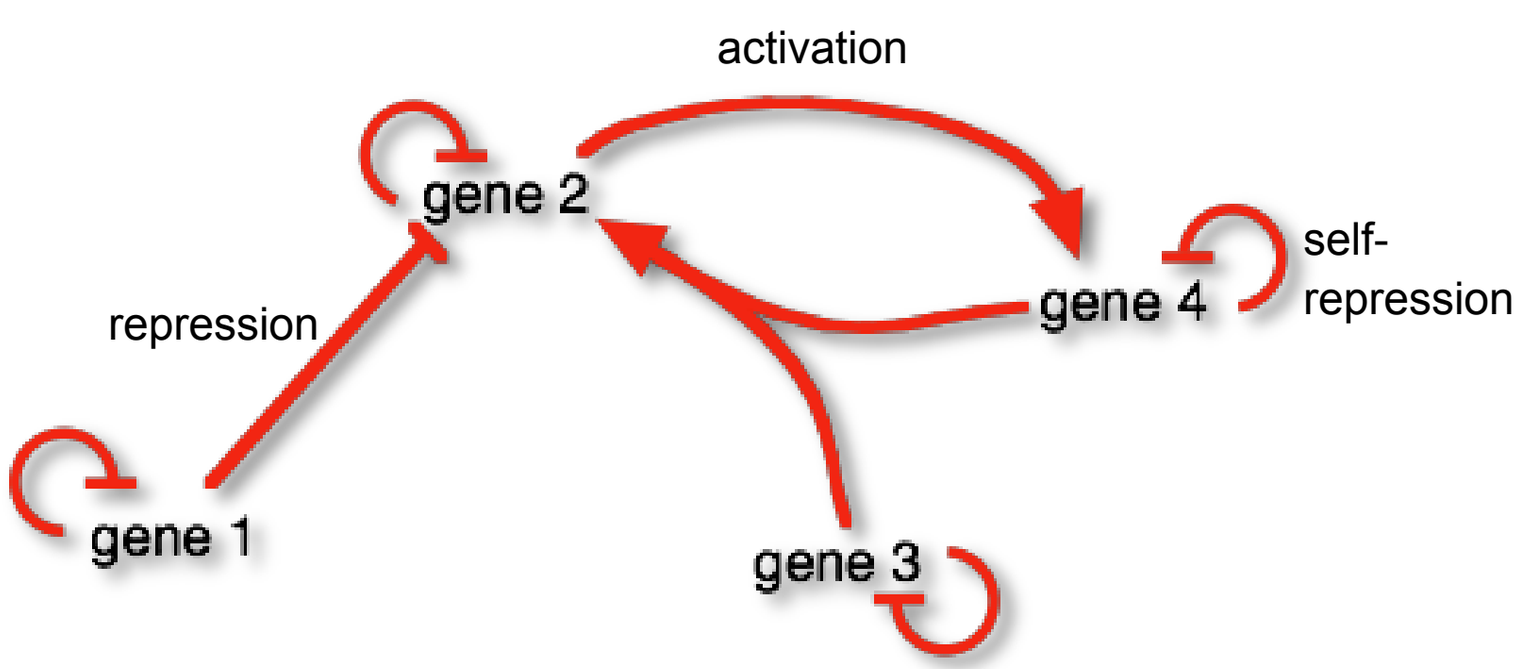
Remember:
genes do not interact directly

Conventions for GRN Graphs

Nodes: genes that code for proteins which catalyze products …
→ everything projected onto respective gene

Gene regulation networks have "cause and action"
→ **directed** networks

A gene can enhance or suppress the expression of another gene
→ **two types** of arrows



Luminescence in *V. fischeri*

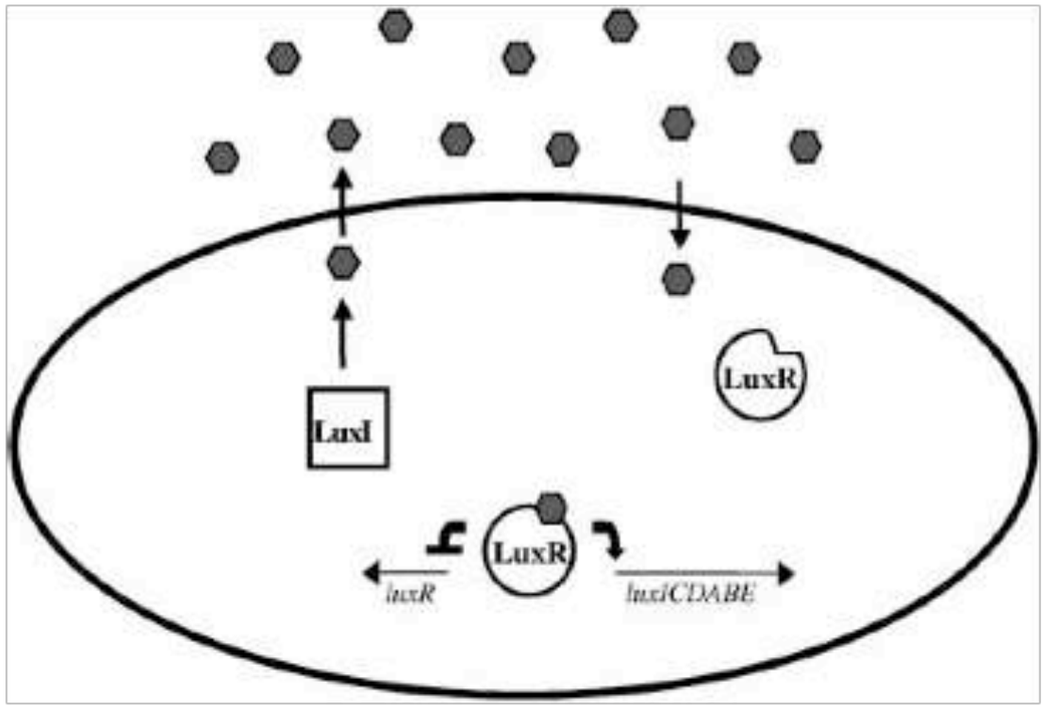


Figure 1 The *Vibrio fischeri* LuxI/LuxR quorum sensing circuit. There are five luciferase structural genes (*luxCDABE*) and two regulatory genes (*luxR* and *luxI*) required for quorum sensing–controlled light emission in *V. fischeri*. The genes are arranged in two adjacent but divergently transcribed units. *luxR* is transcribed to the left, and the *luxICDABE* operon is transcribed to the right. The LuxI protein (square) is responsible for synthesis of the HSL autoinducer *N*-(3-oxohexanoyl)-homoserine lactone (hexagons). As the cell-population density increases, the concentration of the autoinducer increases both intra- and extracellularly. At a critical autoinducer concentration, the LuxR protein (circle) binds the autoinducer. The LuxR-autoinducer complex binds at the *luxICDABE* promoter and activates transcription of this operon. This action results in an exponential increase in autoinducer synthesis via the increase in transcription of *luxI* and an exponential increase in light production via the increase in transcription of *luxCDABE*. The LuxR-autoinducer complex also binds at the *luxR* promoter, but in this case the complex represses the transcription of *luxR*. This negative action compensates for the positive action at the *luxICDABE* promoter. The oval represents a bacterial cell.

Miller, Bassler, 2001

V. fischeri lives in symbiotic association with eukaryotic hosts.

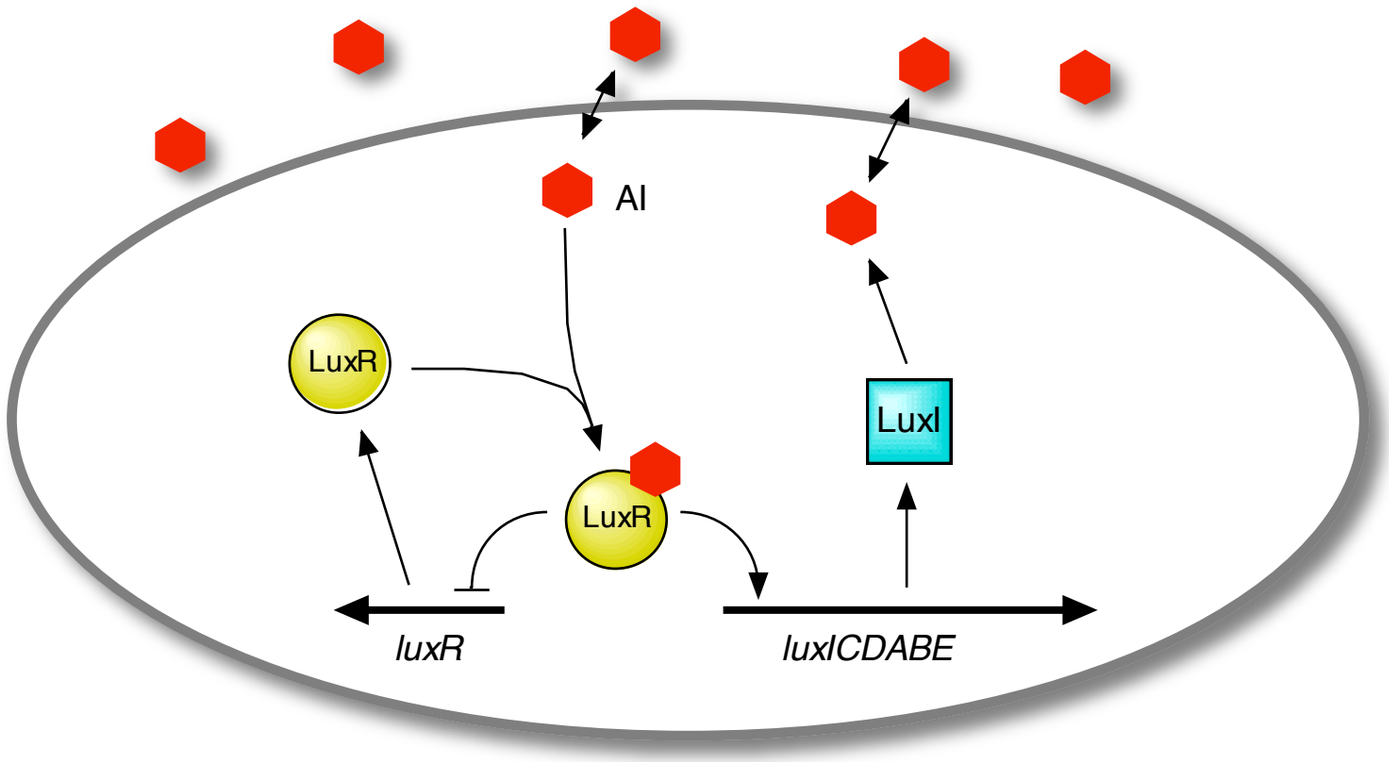
Function: generate light

Squid: camouflage against moon light

Fish *Monocentris japonicus*: attracts a mate by light seduction

The Complete Picture?

Sketch from Miller & Bassler used to explain the mechanism:

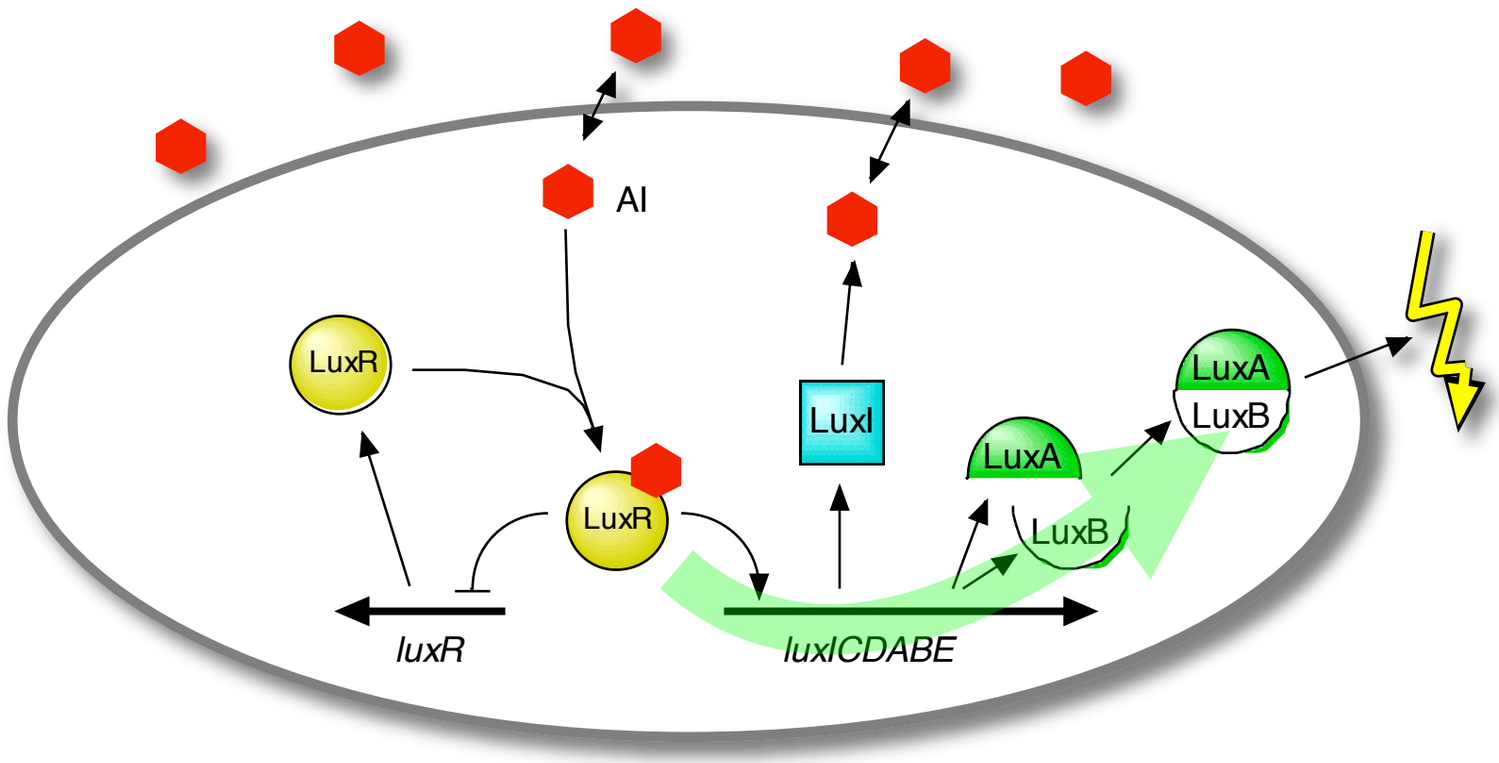


What is **missing**?

- emitted light signal
- degradation of AI and proteins
- threshold
- details of the "reactions"

A Slightly More Complete Picture

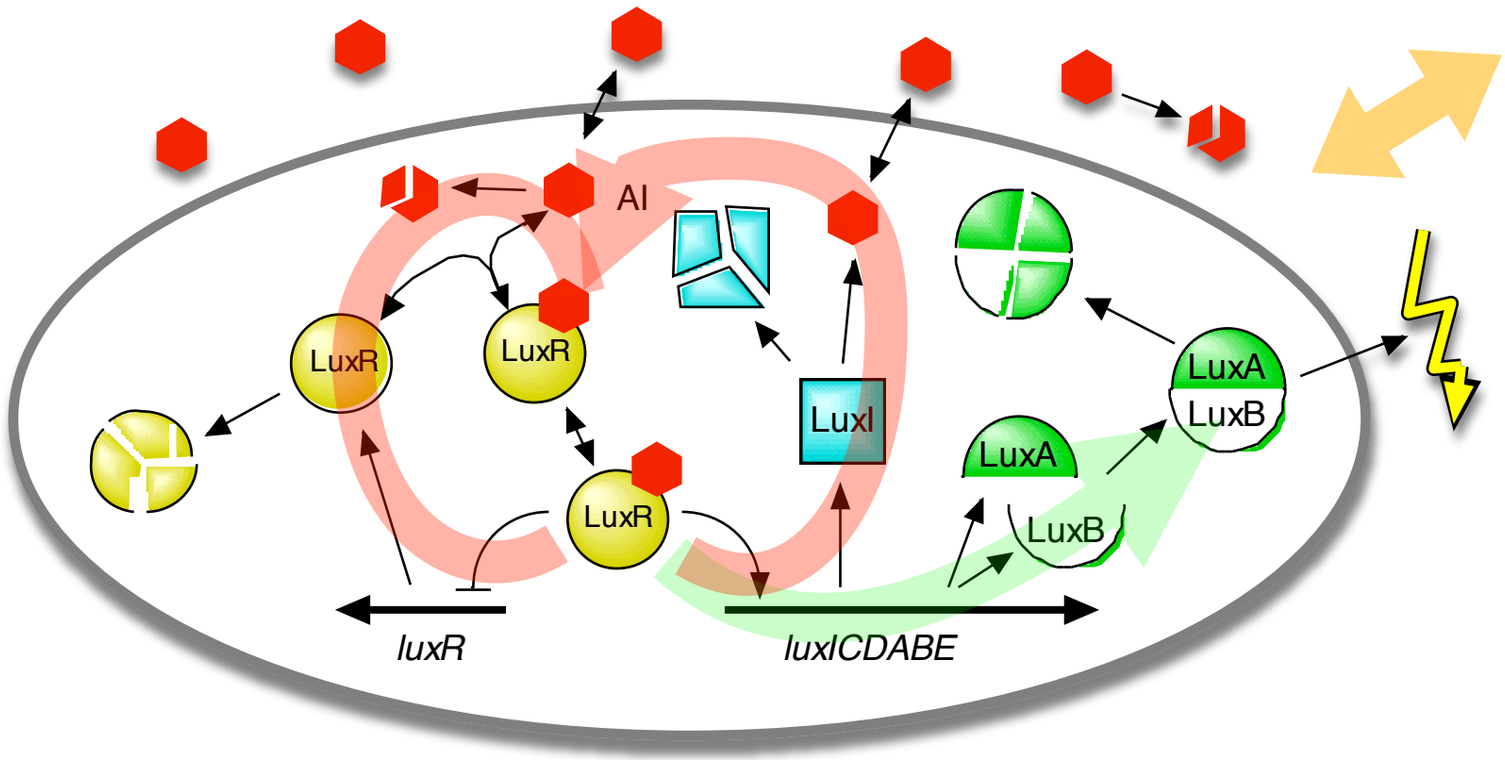
Add luminescence LuxA + LuxB → luciferase → light

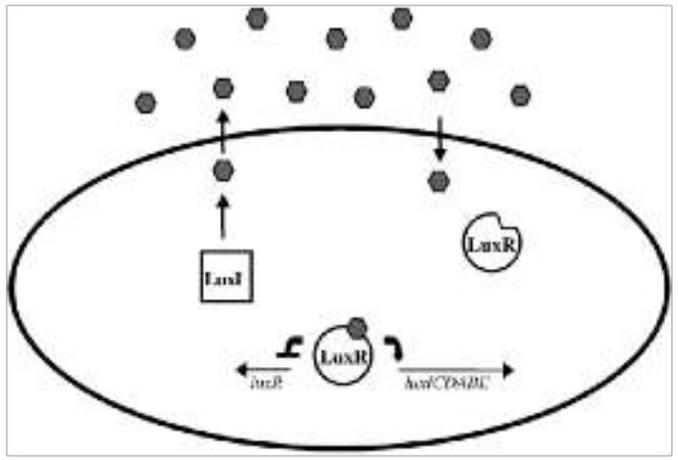


Beware: the picture contains **different reaction mechanisms**
(various associations, transcription + translation, diffusion, …)

Still not the Complete Picture

Add degradation and oligomerisation:





Still missing here:

- dimerization of LuxR:AI
- the Ain-pathway

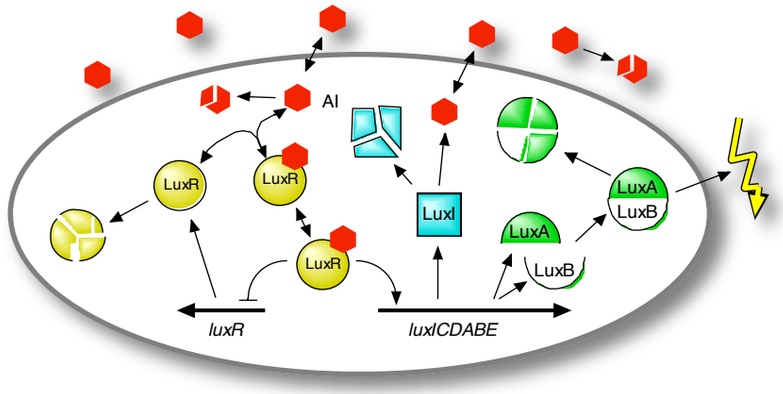
Modeling problem in biology:

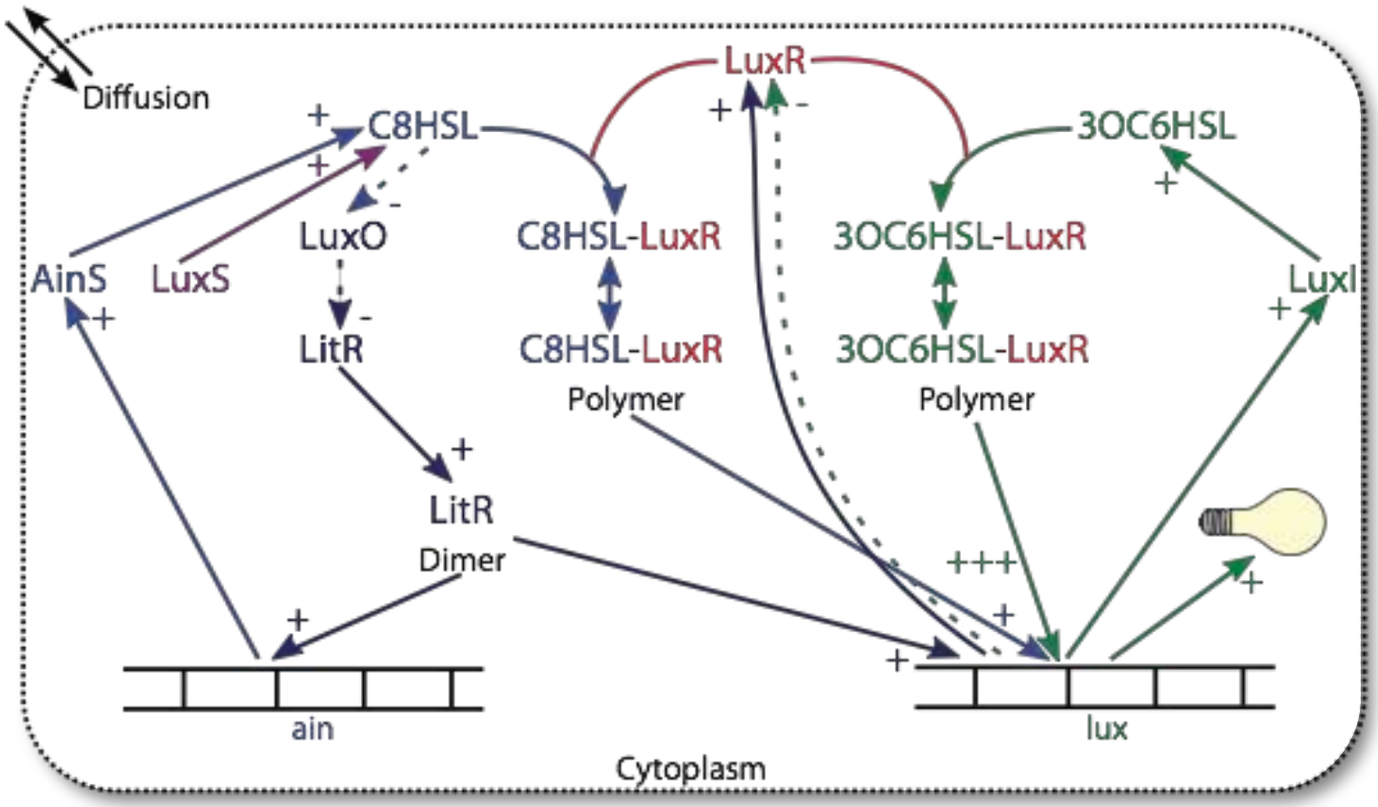
→ convert hand-waving **verbal descriptions** into consistent models

And There is One More Detail…

- two auto-inducers: 3OC6HSL and C8HSL
- two genes (ain and lux)

But: the model is still incomplete





Nadine Schaadt, BSc. thesis

→ which of all these reactions are important for the dynamic behavior of the system?

→ is everything known?

→ systemic model?

→ interactions with cellular environment

→ predictions?

E. coli Regulatory Network

BMC Bioinformatics

BioMed Central

Research article

Open Access

Hierarchical structure and modules in the *Escherichia coli* transcriptional regulatory network revealed by a new top-down approach

Hong-Wu Ma[1], Jan Buer[2,3] and An-Ping Zeng*[1]

Address: [1]Department of Genome Analysis, GBF – German Research Center for Biotechnology, Mascheroder Weg 1, 38124 Braunschweig, Germany, [2]Department of Mucosal Immunity, GBF – German Research Center for Biotechnology, Mascheroder Weg 1, 38124 Braunschweig, Germany and [3]Medical Microbiology and Hospital Hygiene, Medical School Hannover, Carl-Neuberg-Str. 1, 30625 Hannover, Germany

Email: Hong-Wu Ma - hwm@gbf.de; Jan Buer - jab@gbf.de; An-Ping Zeng* - aze@gbf.de

* Corresponding author

Published: 16 December 2004

BMC Bioinformatics 2004, **5**:199 doi:10.1186/1471-2105-5-199

Received: 28 July 2004
Accepted: 16 December 2004

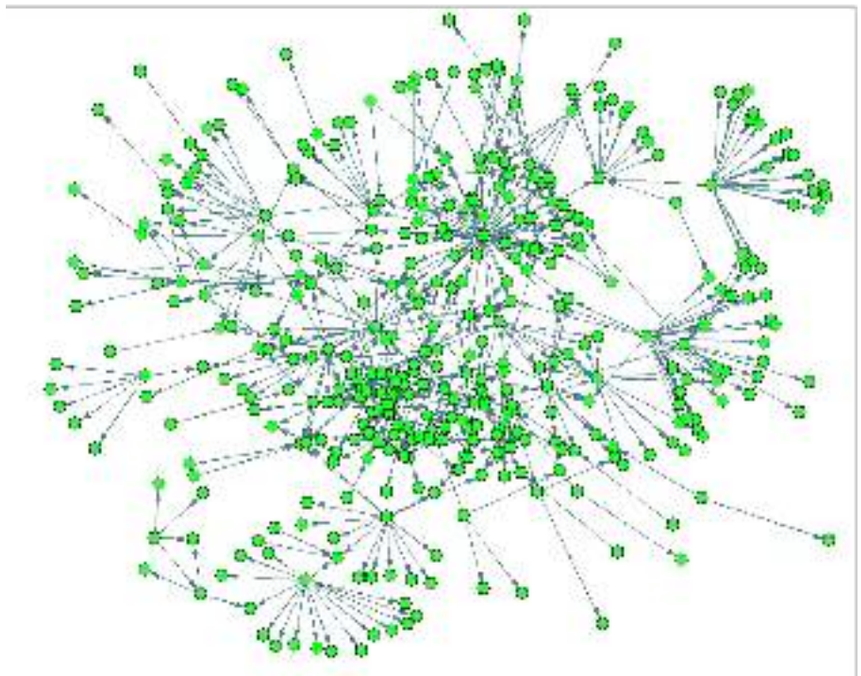
This article is available from: http://www.biomedcentral.com/1471-2105/5/199

BMC Bioinformatics **5** (2004) 199

Hierarchies

Largest WCC: 325 operons (3/4 of the complete network)

WCC = weakly connected component (ignore directions of regulation)



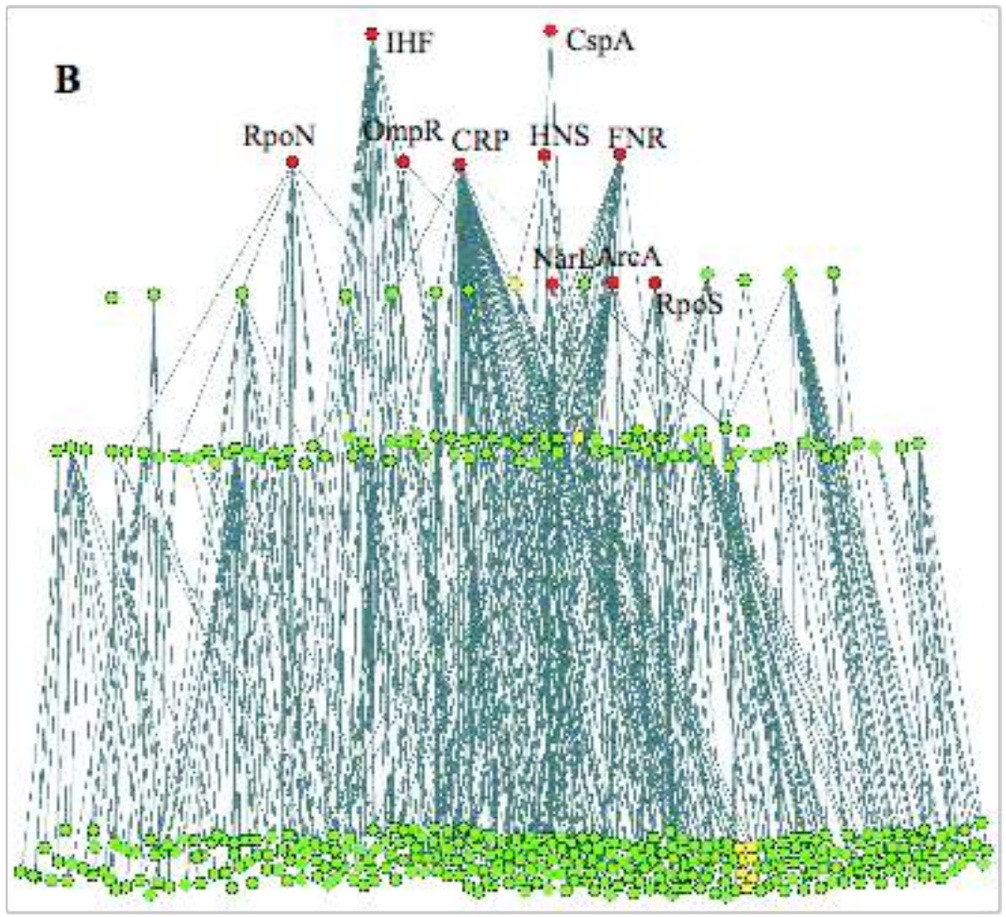
Network from standard layout algorithm

→

Lowest level: operons that code for TFs with only auto-regulation, or no TFs

Next layer: delete nodes of lower layer, identify TFs that do not regulate other operons in this layer (only lower layers)

Continue …



Network with all regulatory edges pointing downwards

→ a few global regulators (•) control all the details

Ma et al., *BMC Bioinformatics* **5** (2004) 199

Global Regulators in *E. coli*

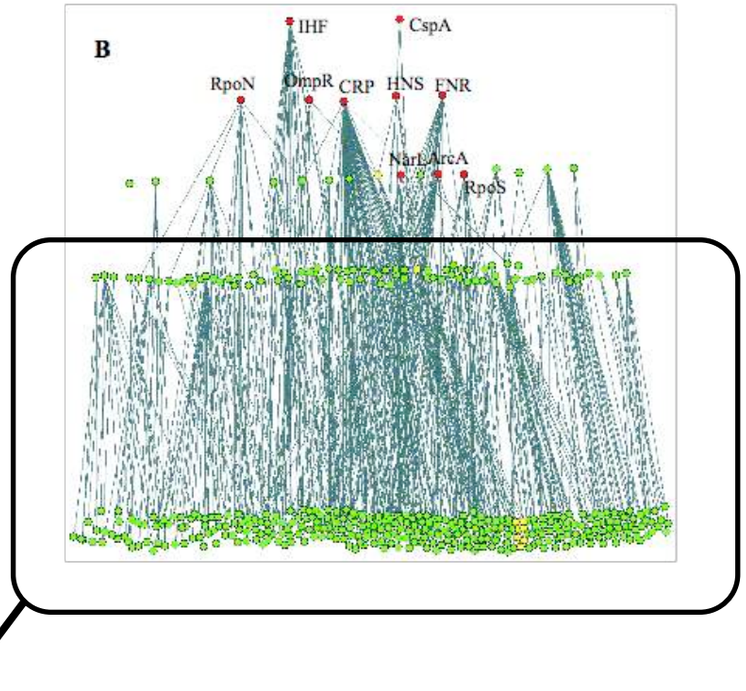
Table 1: Global regulators and their regulated operons and functions in the regulatory network of *E. coli*.

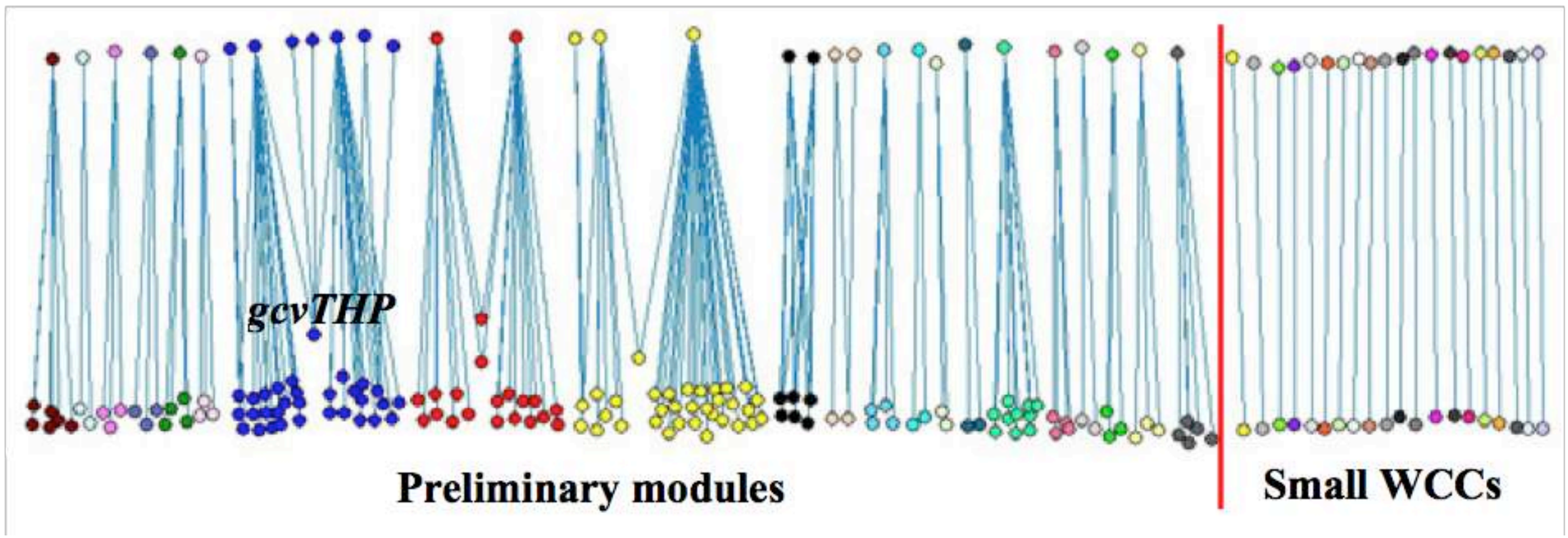
Global regulator	directly regulated Operons	Total regulated operons	Modules regulated	Function
IHF	21	39	15	integration host factor
CspA	2	24	5	Cold shock protein
CRP	72	112	21	cAMP receptor protein
FNR	22	38	16	anaerobic regulator, regulatory gene for nitrite and nitrate reductases, fumarate reductase
HNS	7	22	5	DNA-binding global regulator; involved in chromosome organization; preferentially binds bent DNA
OmpR	6	20	3	Response regulator for osmoregulation; regulates production of membrane proteins
RpoN	12	17	4	RNA polymerase sigma 54 subunit
RpoS	14	24	8	stationary phase sigma factor
ArcA	20	21	6	Response regulator protein represses aerobic genes under anaerobic growth conditions and activates some anaerobic genes
NarL	13	15	5	Two-component regulator protein for nitrate/nitrite response

Ma et al., *BMC Bioinformatics* **5** (2004) 199

Modules

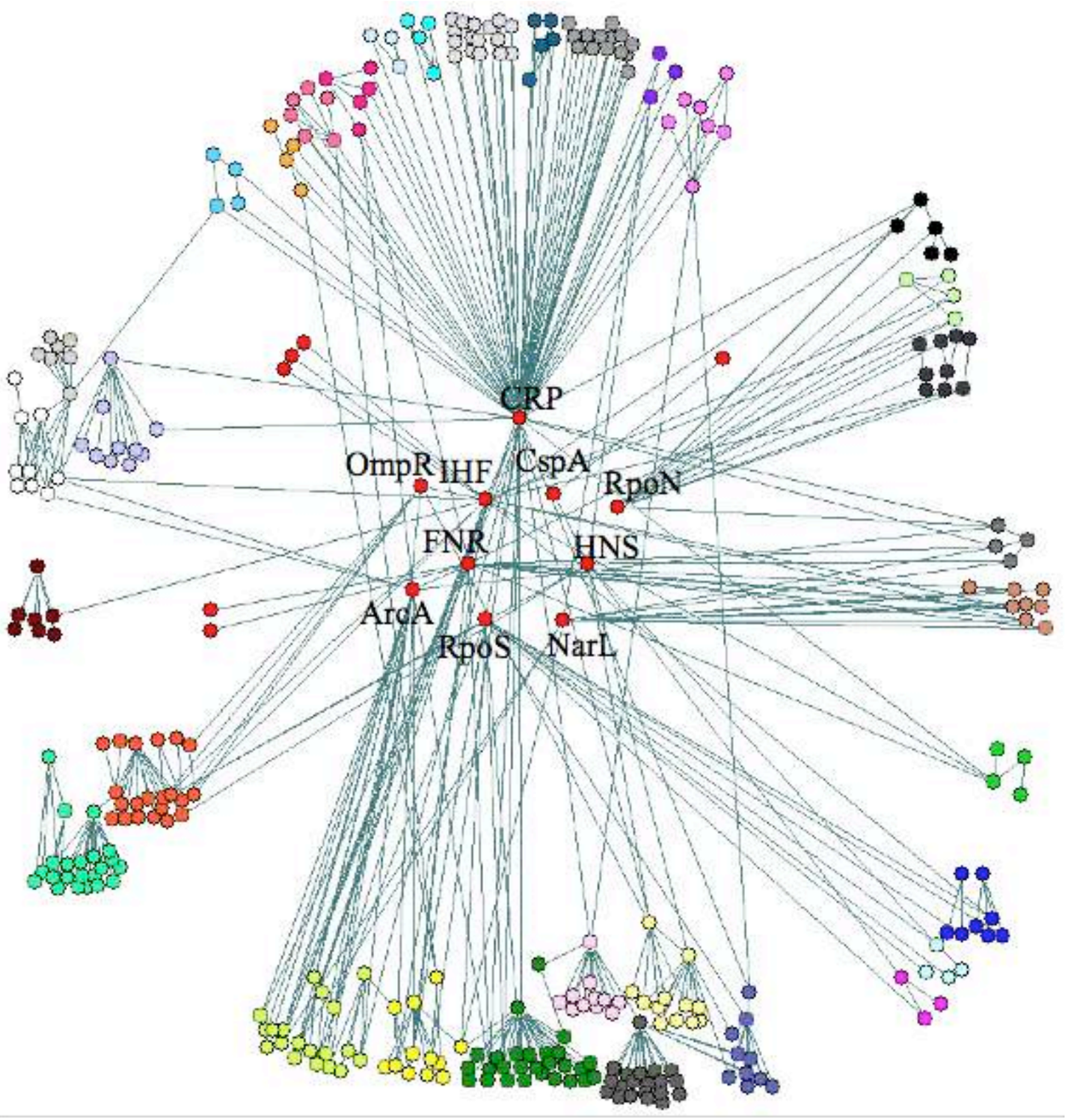
Remove top three layers and determine WCCs
→ just a few modules





Ma et al., *BMC Bioinformatics* **5** (2004) 199

Putting it back together



The ten global regulators are at the core of the network,

some hierarchies exist between the modules

Ma et al., *BMC Bioinformatics* **5** (2004) 199

Naming a few!

Table 2: Functional investigation of modules identified.

index	Operons included	Biological function description
1	*aceBAK, acs, adhE, fruBKA, fruR, icdA, iclMR, mlc, ppsA, ptsG, ptsHI_crr, pykF*	Hexose PTS transport system, PEP generation, Acetate usage, glyoxylate shunt
2	*acnA, fpr, fumC, marRAB, nfo, sodA, soxR, soxS, zwf*	Oxidative stress response
3	*ada_alkB, aidB, alkA, ahpCF, dps, gorA, katG, oxyR*	Oxidative stress response, Alkylation
4	*alaWX, aldB, argU, argW, argX_hisR_leuT_proM, aspV, dnaA, leuQPV, leuX, lysT_valT_lysW, metT_leuW_glnUW_metU_glnVX, metY_yhbC_nusA_infB, nrdAB, pdhR_aceEF_lpdA, pheU, pheV, proK, proL, proP, sdhCDAB_b0725_sucABCD, serT, serX, thrU_tyrU_glyT_thrT, thrW, tyrTV, valUXY_lysV, yhdG_fis*	rRNA, tRNA genes, DNA synthesis system, pyruvate dehydrogenase and ketoglutarate dehydrogenase system
5	*araBAD, araC, araE, araFGH, araJ*	Arabinose uptake and usage
6	*argCBH, argD, argE, argF, argI, argR, carAB*	Arginine usage, urea cycle
7	*caiF, caiTABCDE, fixABCX*	Carnitine usage
8	*clpP, dnaKJ, grpE, hflB, htpG, htpY, ibpAB, lon, mopA, mopB, rpoH*	Heat shock response
9	*codBA, cvpA_purF_ubiX, glnB, glyA, guaBA, metA, metH, metR, prsA, purC, purEK, purHD, purL, purMN, purR, pyrC, pyrD, speA, ycfC_purB, metC, metF, metJ*	Purine synthesis, purine and pyrimidine salvage pathway, methionine synthesis
10	*cpxAR, cpxP, dsbA, ecfI, htrA, motABcheAW, ppiA, skp_lpxDA_fabZ, tsr, xprB_dsbC_recJ*	Stress response, Conjugative plasmid expression, cell motility and Chemotaxis
11	*dctA, dcuB_fumB, frdABCD, yjdHG*	C4 dicarboxylate uptake
12	*edd_eda, gntKU, gntR, gntT*	Gluconate usage, ED pathway
13	*csgBA, csgDEFG, envY_ompT, evgA, gcvA, gcvR, gcvTHP, gltBDF, ilvIH, kbl_tdh, livJ, livKHMGF, lrp, lysU, ompC, ompF, oppABCDF, osmC, sdaA, serA, stpA*	Amino acid uptake and usage
14	*fdhF, fhlA, hycABCDEFGH, hypABCDE*	Formate hydrogenlyase system
15	*flgAMN, flgBCDEFGHIJ, flgKL, flgMN, flhBAE, flhDC, fliAZY, fliC, fliDST, fliE, fliFGHIJK, fliLMNOPQR, tarTapcheRBYZ*	Flagella motility system
16	*ftsQAZ, rcsAB, wza_wzb_b2060_wcaA_wcaB*	Capsule synthesis, cell division
17	*gdhA, glnALG, glnHPQ, nac, putAP*	Glutamine and proline utilization
18	*glmUS, manXYZ, nagBACD, nagE*	Glucosamine, mannose utilization
19	*glpACB, glpD, glpFK, glpR, glpTQ*	Glycerol phosphate utilization
20	*lysA, lysR, tdcABCDEFG, tdcR*	Serine, threonine usage
21	*malEFG, malK, lamB, malM, malPQ, malS, malT, malZ*	Maltose utilization

Ma et al., *BMC Bioinformatics* **5** (2004) 199

Summary

- **Static** PPI networks:
  → topology, measures, data sources, …

- **Changes** during cell cycle, adaptation to environmental changes, …
  → Gene Regulation
    → many biological steps
      → often modeled on the gene level only

Next lecture:

- Regulatory **motifs**
  → static and dynamic behavior