Networks, noise and survival in stress

Gábor Balázsi
Department of Systems Biology – U950
The University of Texas M. D. Anderson Cancer Center
Houston, TX 77054, USA

The Microsoft Research - University of Trento
Centre for Computational and Systems Biology
Computational and Systems Biology Course at CoSBi
Wednesday, February 19, 2008

THE UNIVERSITY OF TEXAS
MD ANDERSON
CANCER CENTER
Making Cancer History®

Networks (Graphs)

- A ***network (graph)*** is a system of interconnected objects

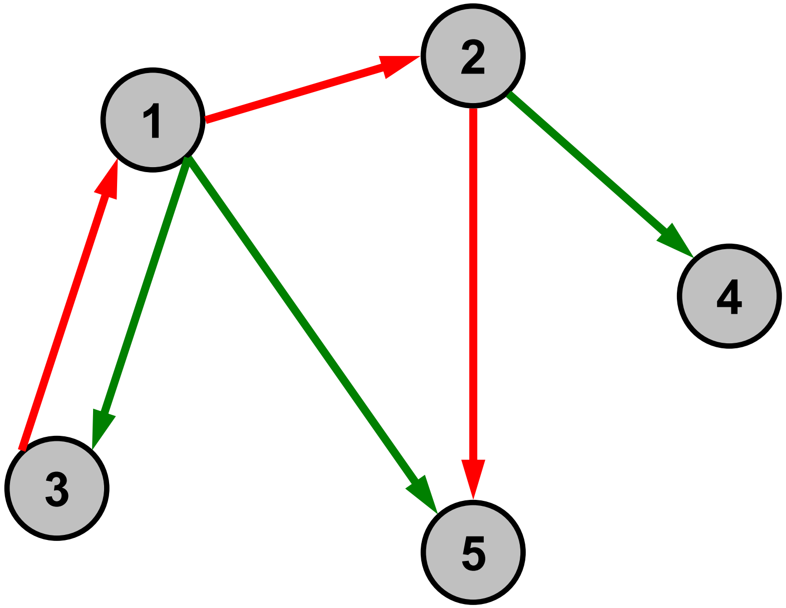
- Components of a network:

  – ***Nodes (vertices)***

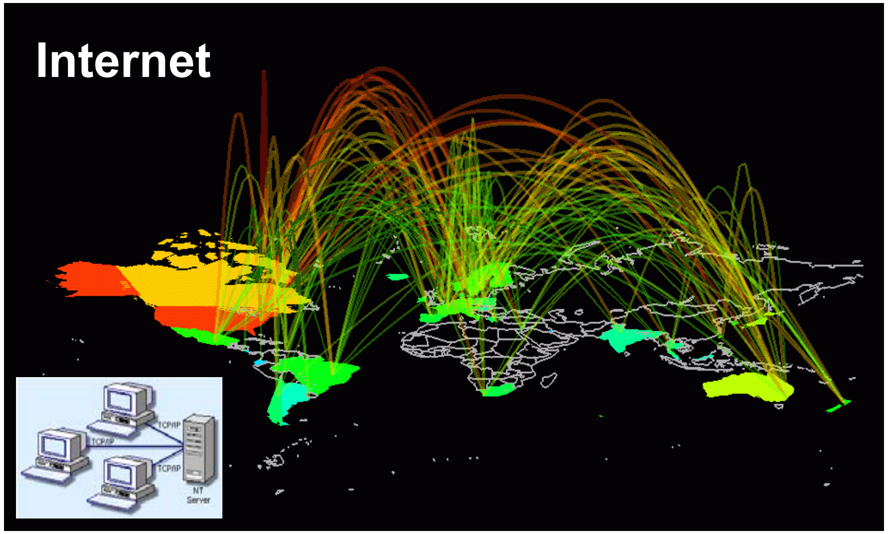
  – ***Links (edges) can be:***

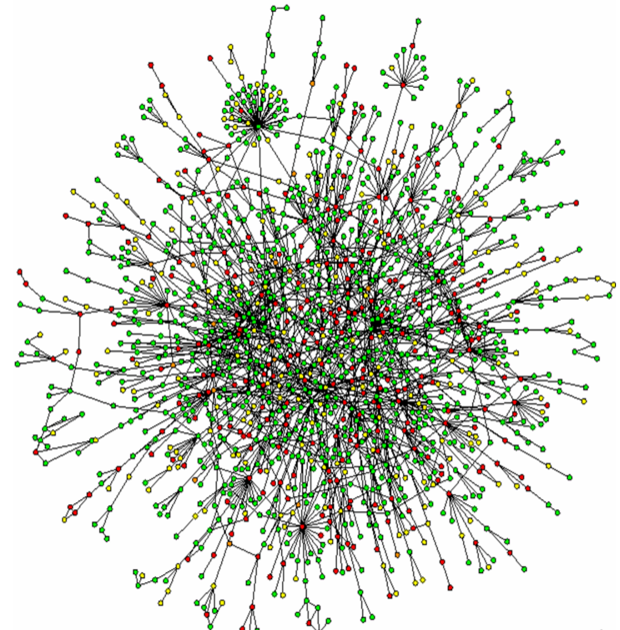
    - Directed

    - Undirected



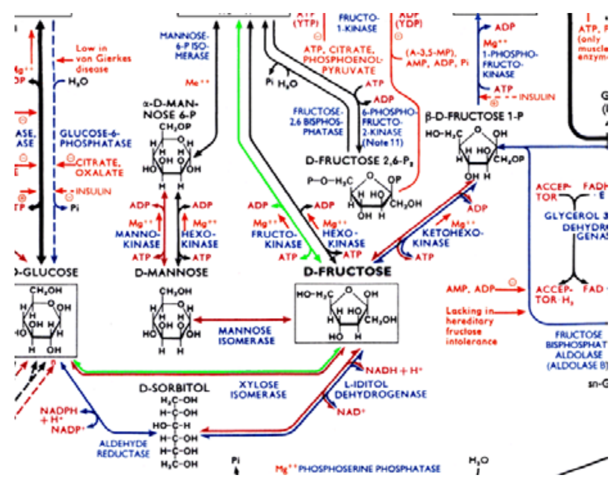
Examples of networks





Protein-protein interaction network

Social network

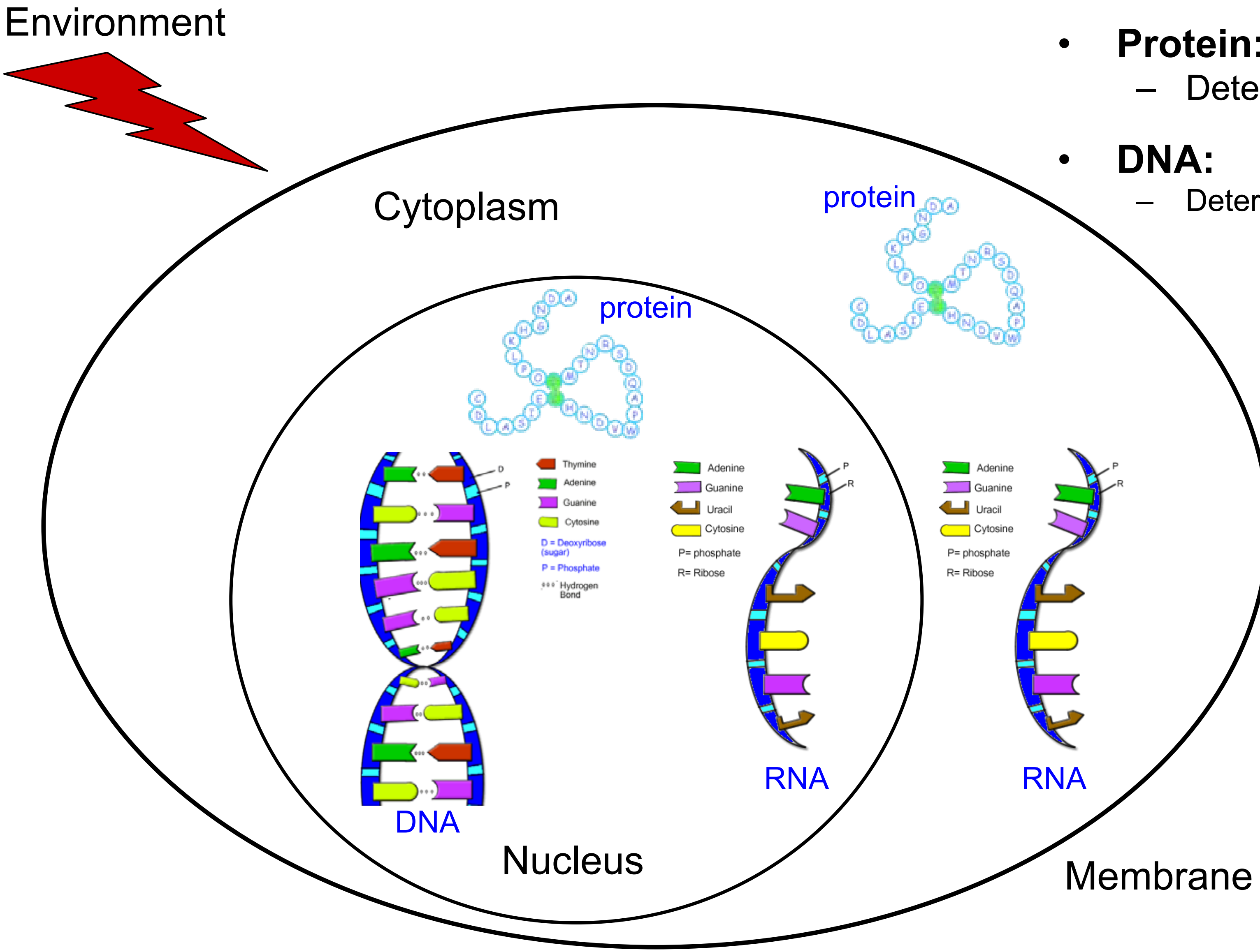


Metabolic network

Network topology: A large-scale perspective (Part I)

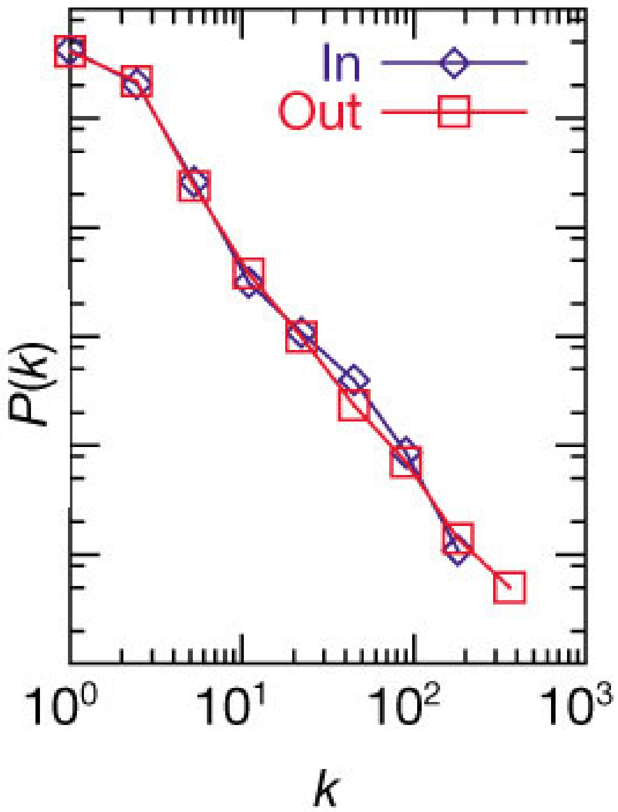
Gene expression

- **Protein:**
  - Determines phenotype
- **DNA:**
  - Determines genotype

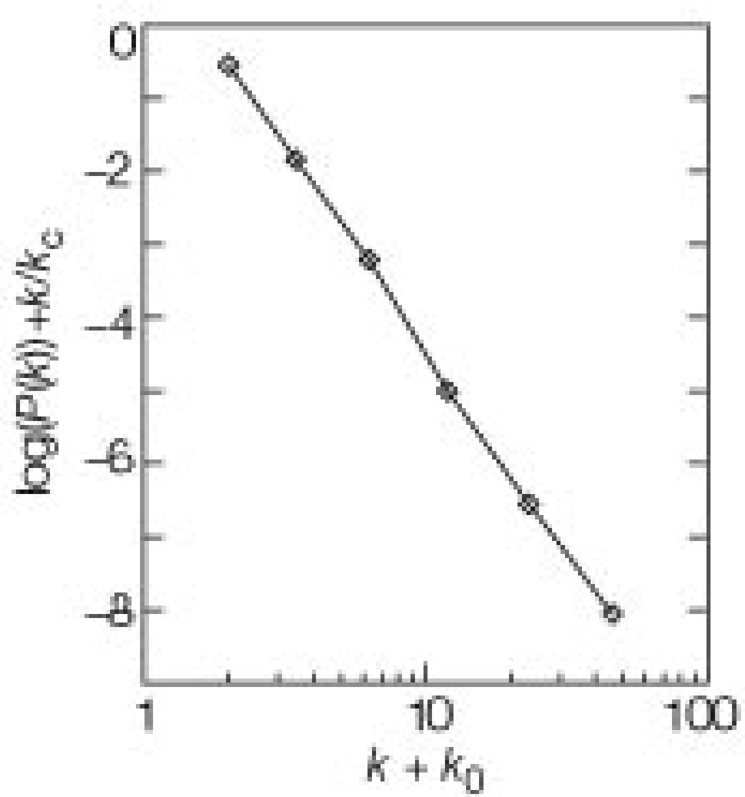


Connectivity distribution

METABOLIC



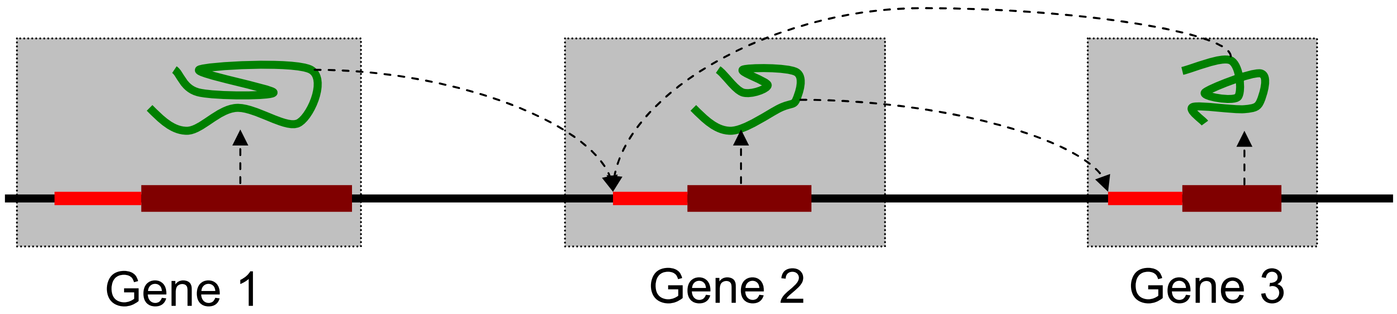
PPI

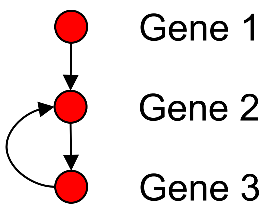


Jeong et al., *Nature* **407**, 651 (2000);

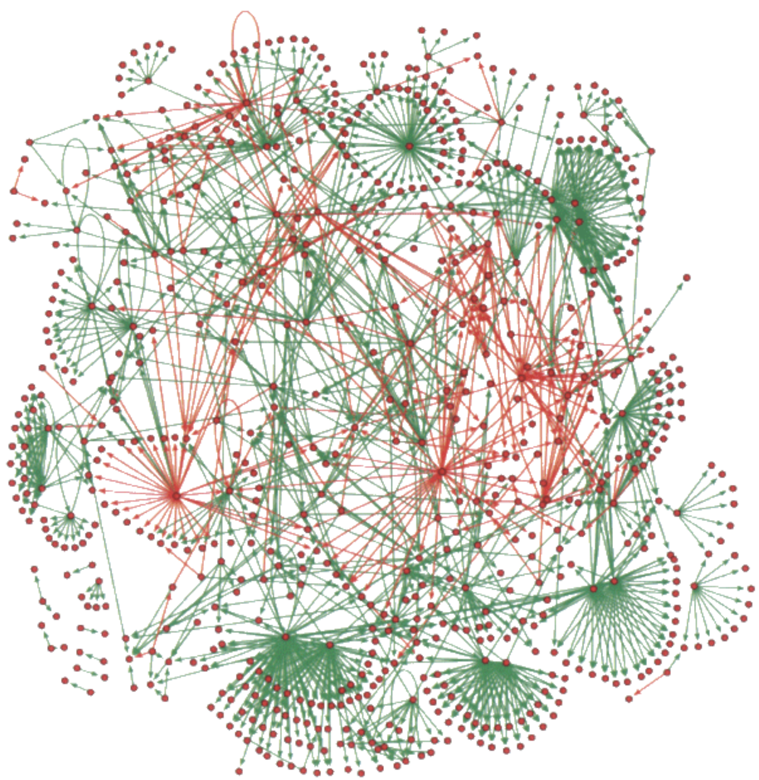
Jeong et al., *Nature* **411**, 41 (2001)

Transcriptional regulatory (TR) networks

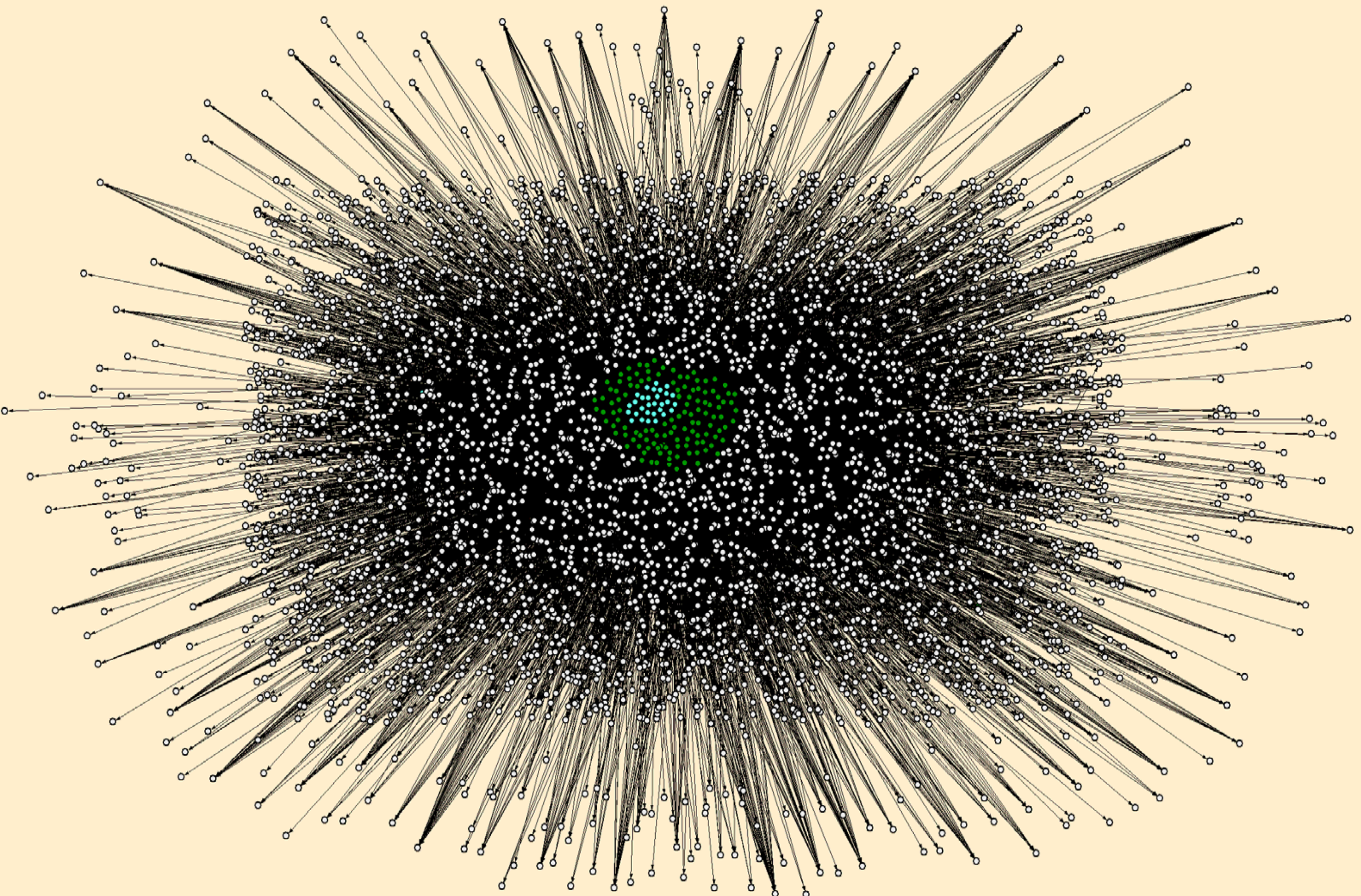




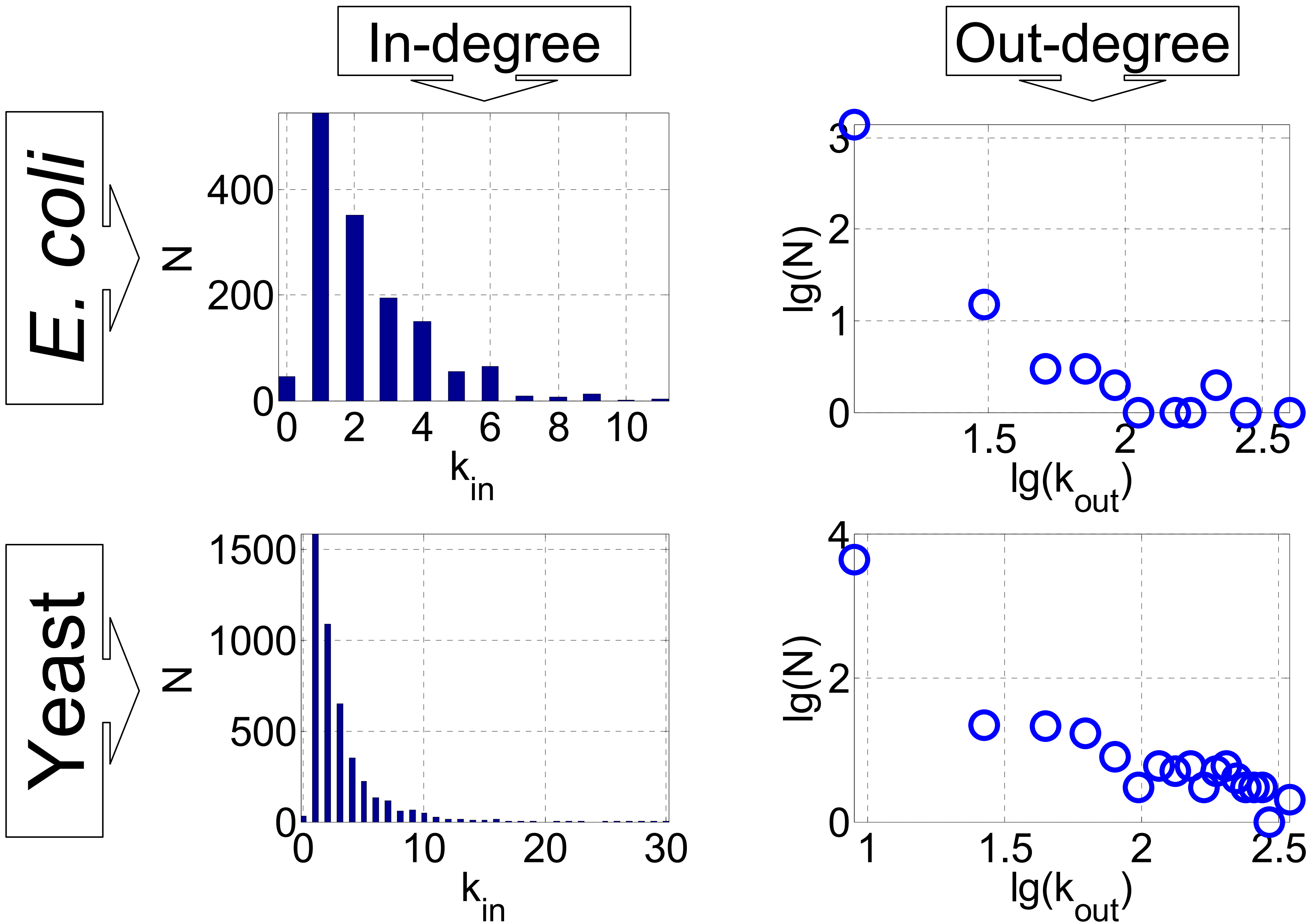
Maslov & Sneppen, Phys. Biol. 2005



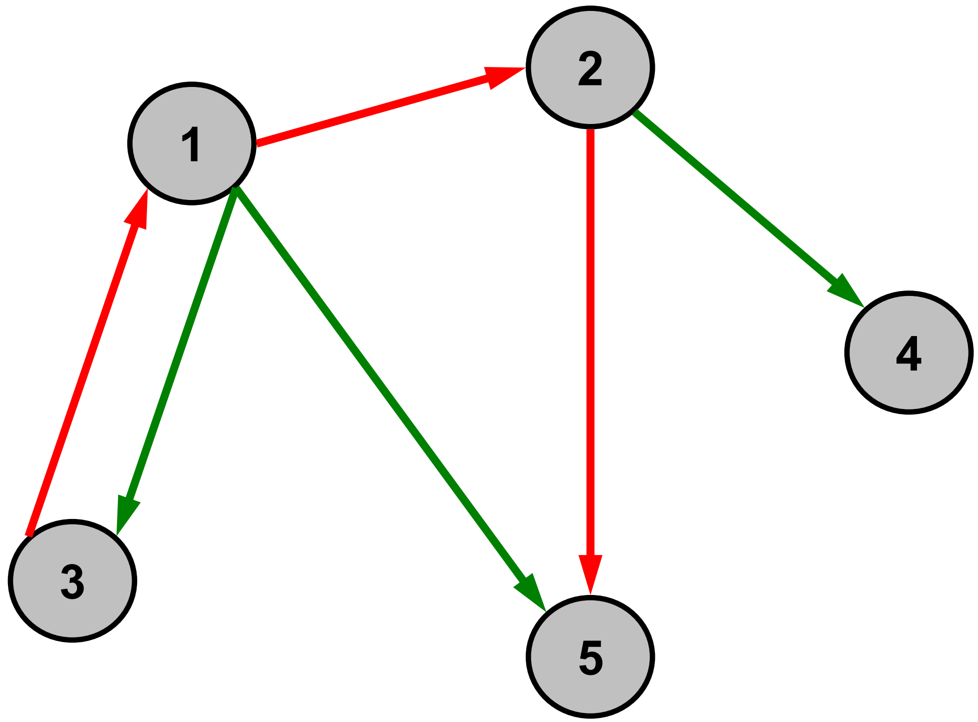
The yeast TR network



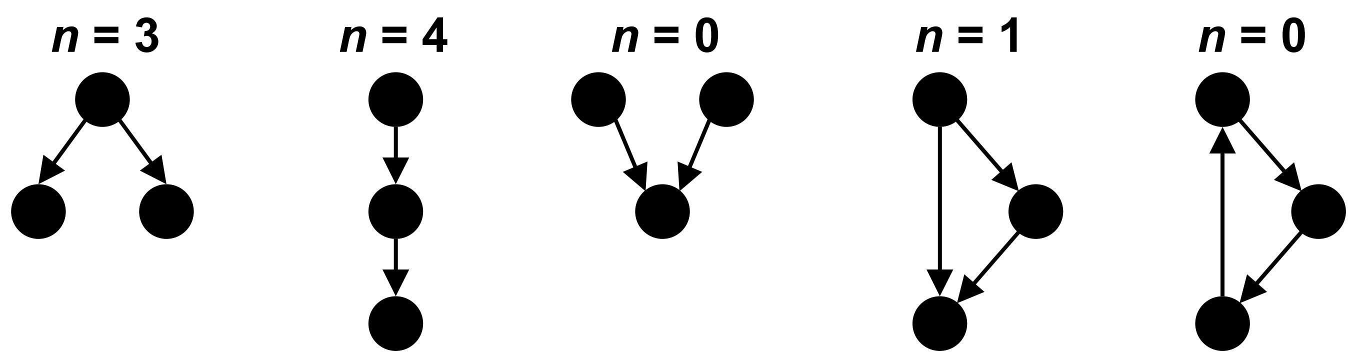
Connectivity distribution of TR networks



Subgraphs

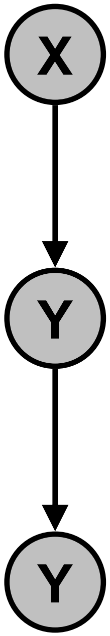


Abundance of 3-node subgraphs

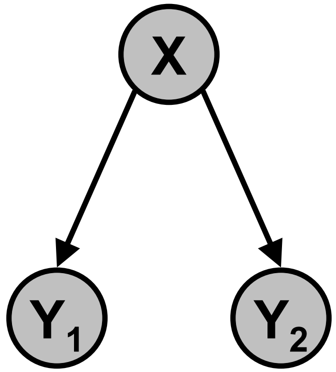


Subgraphs: some examples

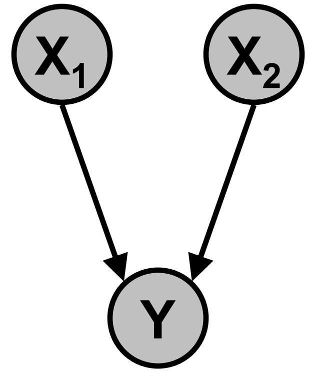
CAS
(Cascade)



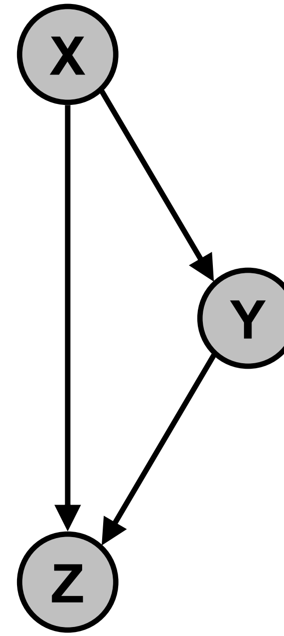
DIV
(Divergence)



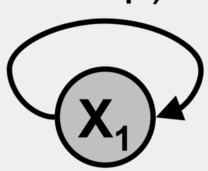
CNV
(Convergence)



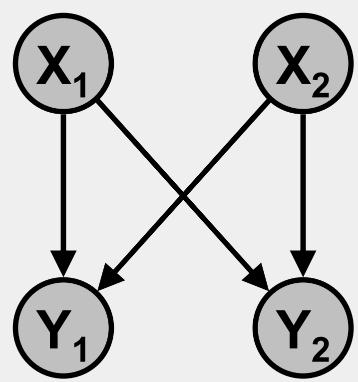
FFL
(Feed-forward loop)



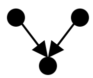
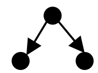
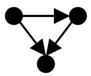
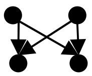
ARL
(Autoregulatory loop)



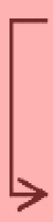
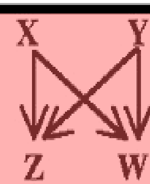
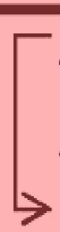
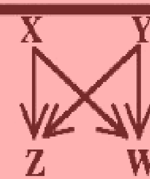
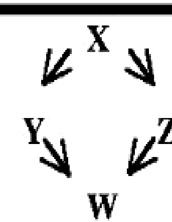
BFM
(Bi-fan motif)



Subgraph abundance in the E. coli TR network

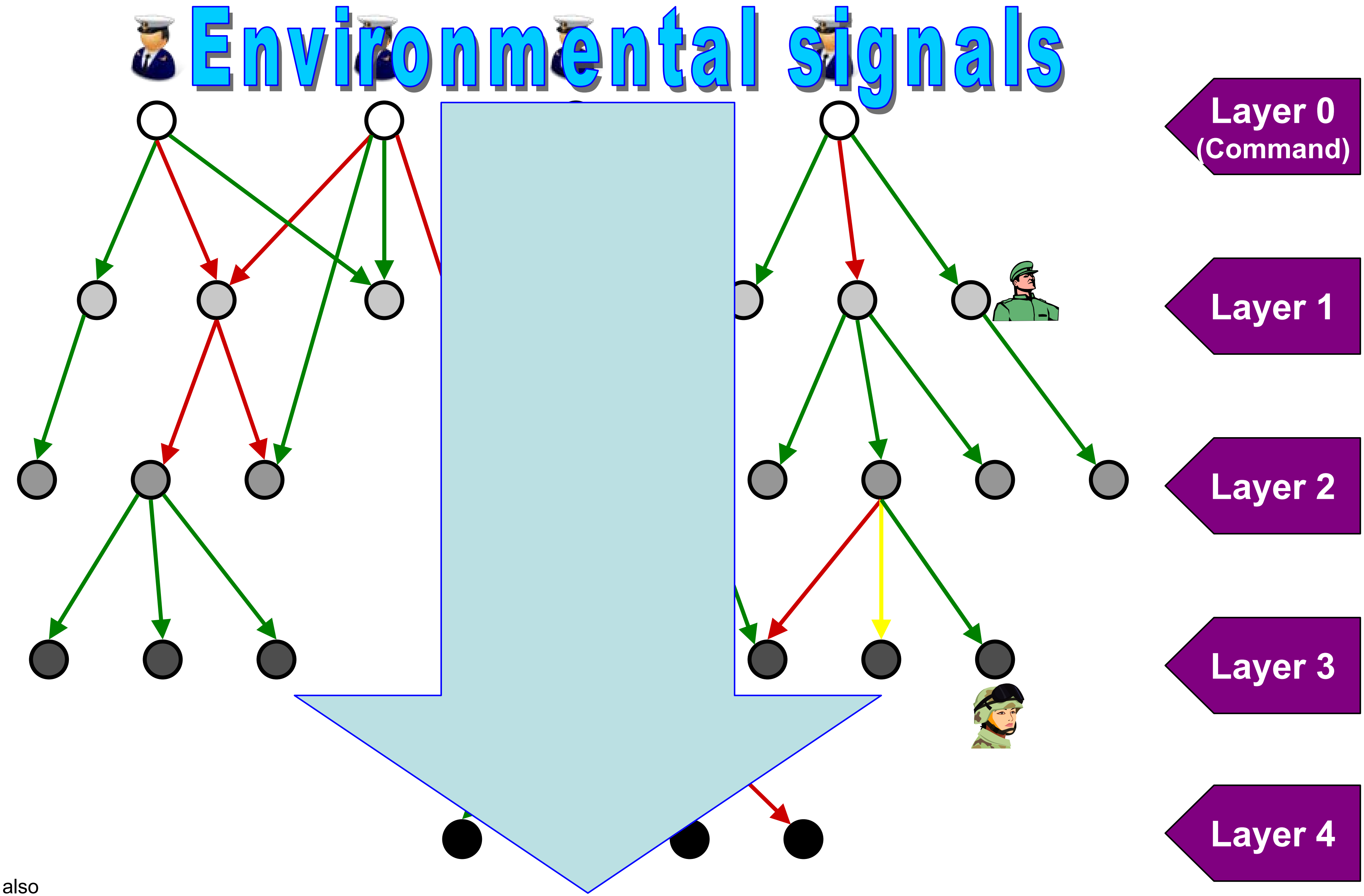
Subgraph	**CON** 	**DIV** 	**CAS** 	**FFL** 	**BFM** 
Abundance in *E. coli* network	227	4777	160	42	209
Abundance in randomized network	231.92 ± 8.05	4339.3 ± 132.0	186.69 ± 7.08	9.50 ± 4.17	74.78 ± 16.36
Z-score	0.6114	3.3149	3.7682	7.7909	8.2052

Motifs in information-processing networks

Network	Nodes	Edges	N_{real}	$N_{rand} \pm SD$	Z score	N_{real}	$N_{rand} \pm SD$	Z score	N_{real}	$N_{rand} \pm SD$	Z score
Gene regulation (transcription)					**Feed-forward loop**			**Bi-fan**			
E. coli	424	519	40	7 ± 3	10	203	47 ± 12	13			
*S. cerevisiae**	685	1,052	70	11 ± 4	14	1812	300 ± 40	41			
Neurons					**Feed-forward loop**			**Bi-fan**			**Bi-parallel**
C. elegans†	252	509	125	90 ± 10	3.7	127	55 ± 13	5.3	227	35 ± 10	20

R Milo et al., *Science* **298**, 824-827 (2002).

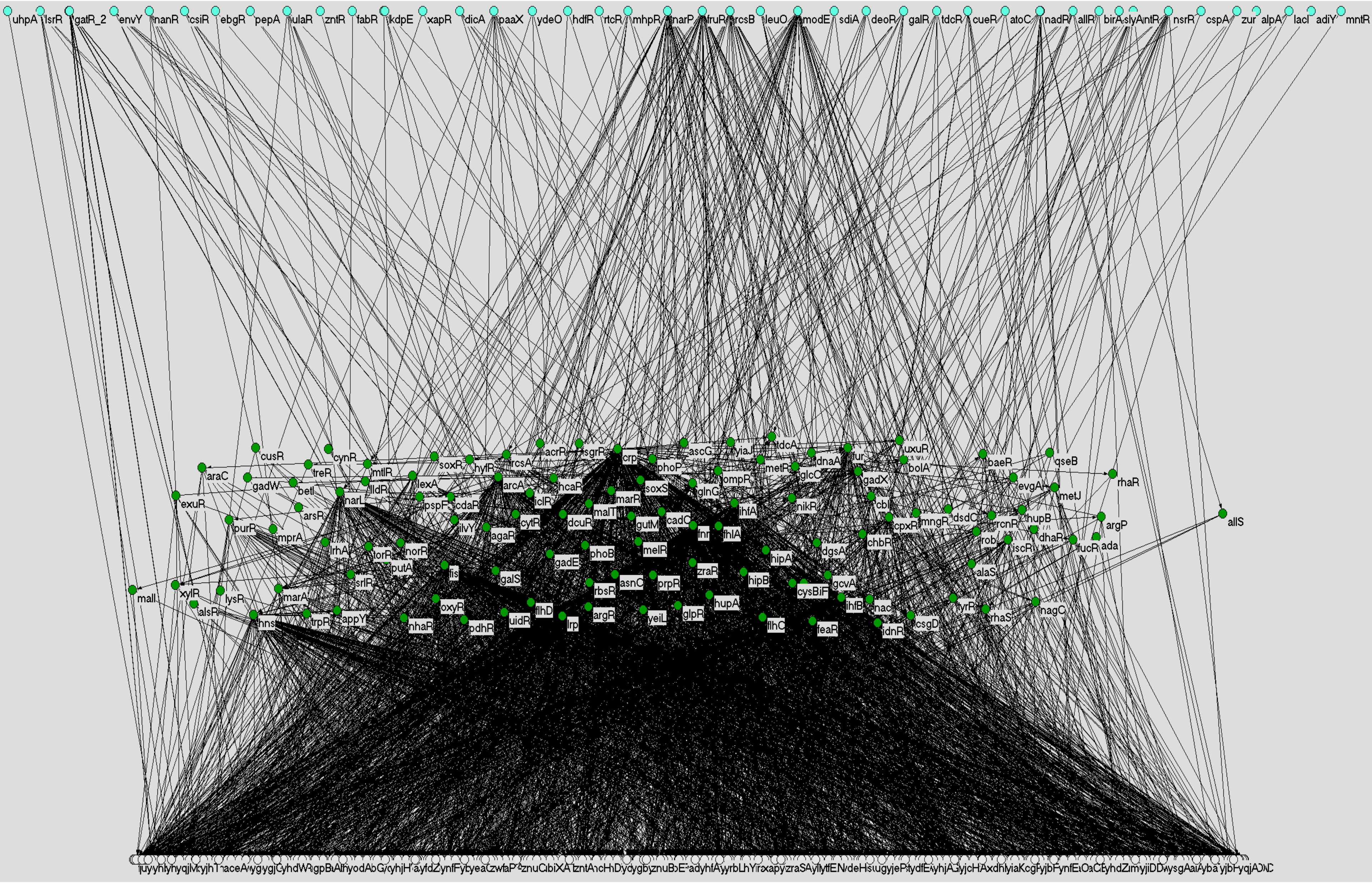
Hierarchical TR network topology



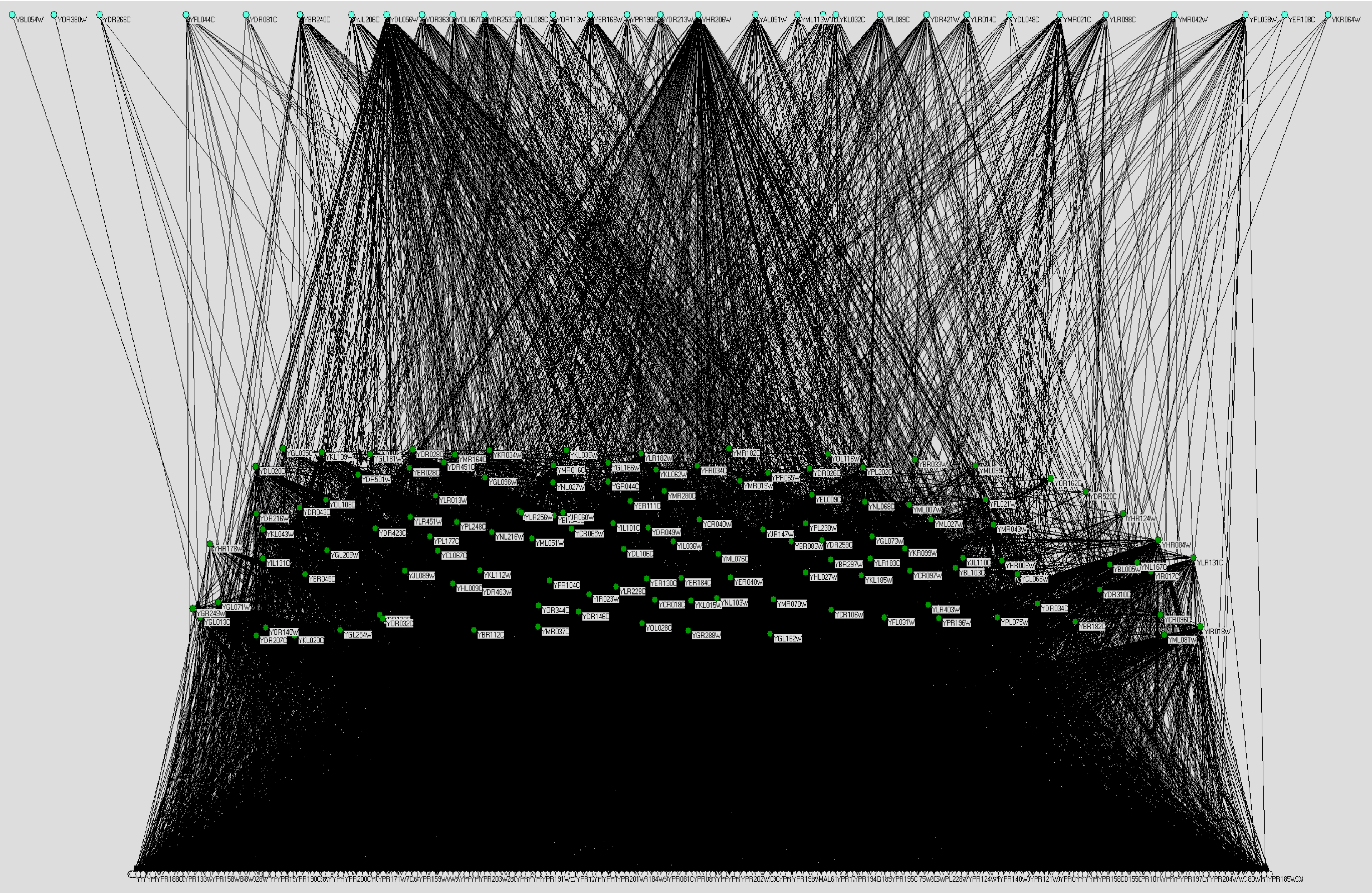
See also
Ma et al., *BMC Bioinformatics* **5**, 199 (2004)
Yu and Gerstein, *PNAS* **103**, 14724 (2006).

Balázsi G, Barabási A.-L., Oltvai ZN, *PNAS* **102**, 7841-6 (2005)

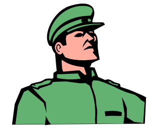
The *E. coli* TR network



The yeast TR network



Types of genes: Commander, Intermediate, Executor


82 + 78 + 1273 = 1433	Commander	Intermediate	Executor	Total
Number of genes	**82**	**78**	**1273**	**1433**

Escherichia coli

31 + 126 + 4284 = 4441	Commander	Intermediate	Executor	Total
Number of genes	**31**	**126**	**4284**	**4441**

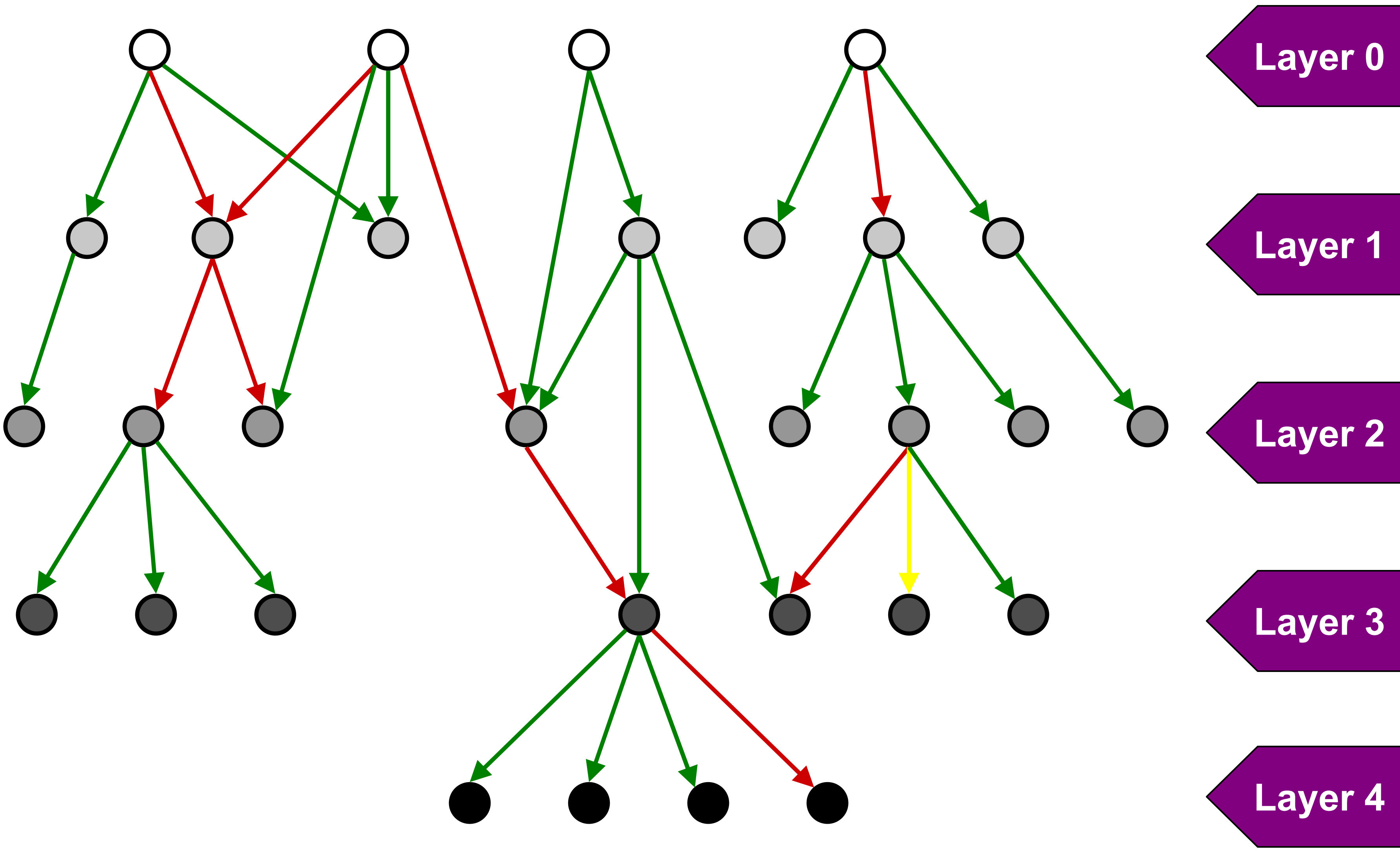
Saccharomyces cerevisiae

34 + 11 + 735 = 780	Commander	Intermediate	Executor	Total
Number of genes	**34**	**11**	**735**	**780**

Mycobacterium tuberculosis

Are all three gene types equally responsive and/or noisy?

Definition of origons



See also Ma'ayan et al., *Science* **309**, 1078 (2005) & Balázsi et al., *PNAS* **102**, 7841 (2005)

Origons in the E. coli regulatory network

Why are origons important?

- Many nodes serve as ***sensors***, monitoring environmental changes. The information captured by the sensor TFs percolate into their origons.

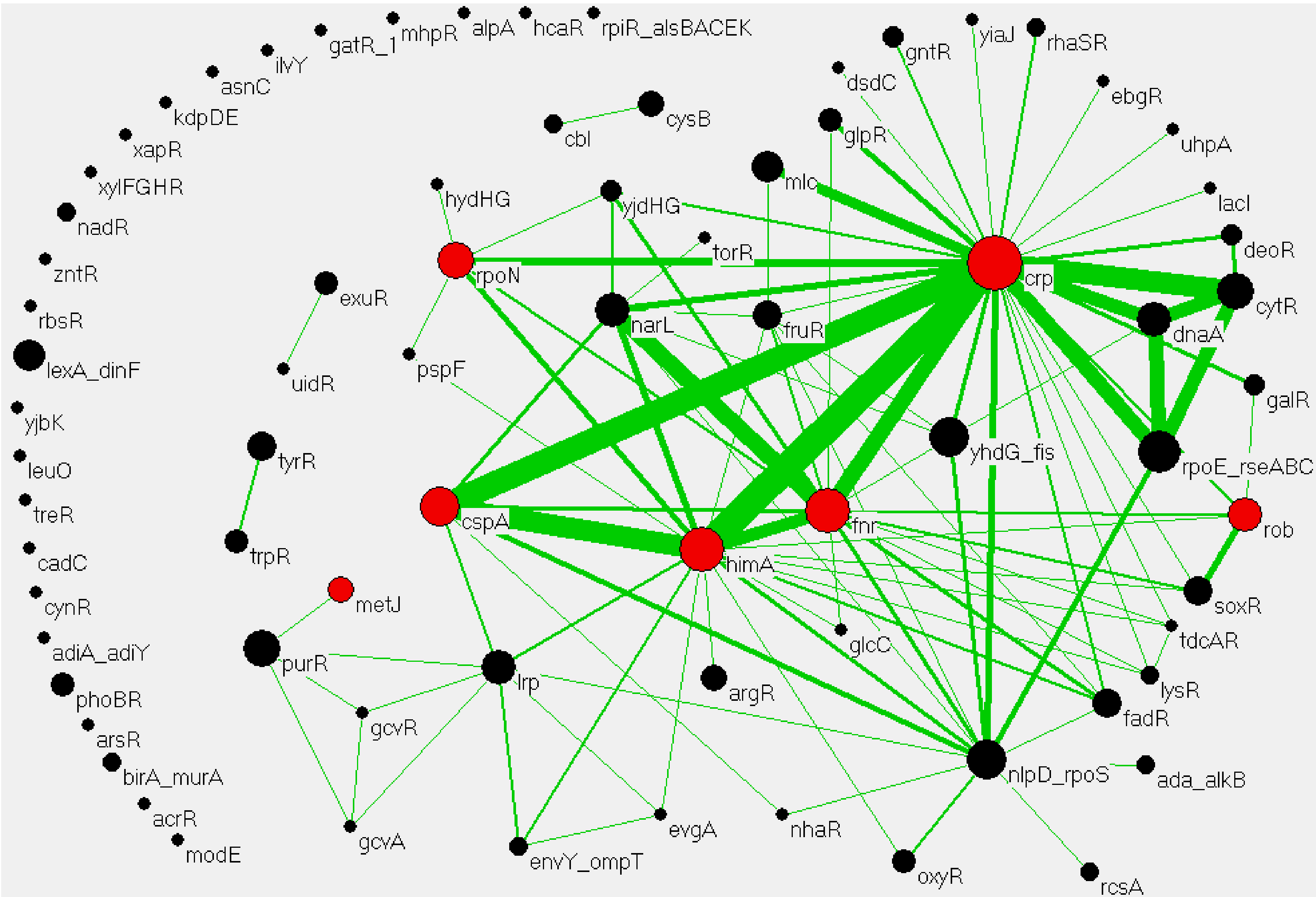
- If we suppose it is possible to perturb only one node in the *E. coli* TR network (i.e., by altering input gene expression or the activity of a transcription factor), primarily the nodes ***within the origon*** are expected to change their expression levels due to transcriptional regulation.

- If most of the network is unaffected, ***only the origon of the perturbed input node*** remains to be analyzed, ***reducing network complexity***.

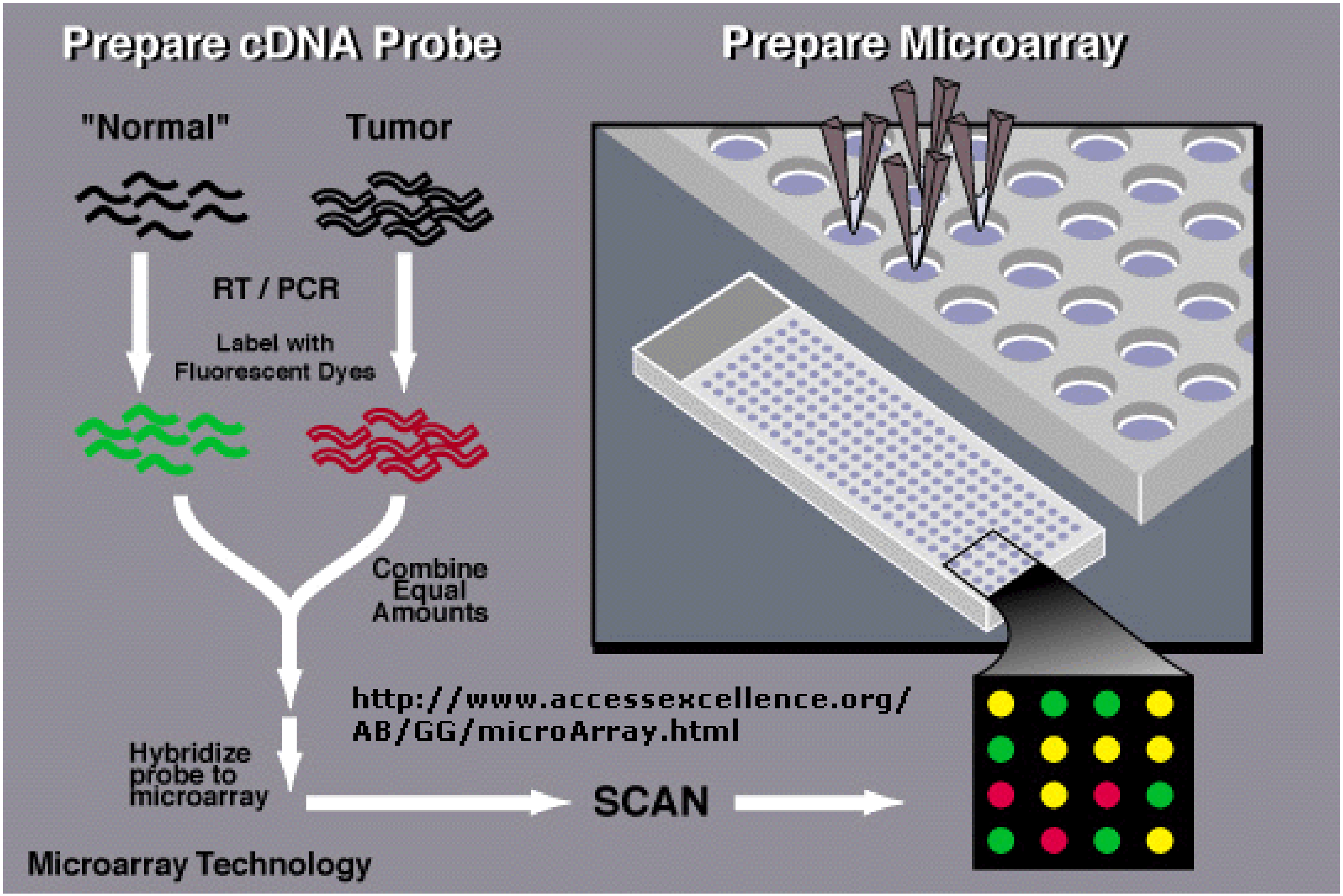
- Origons are more complex than **modulons**, but less complex than **stimulons**

- If **specific** origons respond to **specific** stimuli, then origons are the **topological units of dynamical network utilization**.

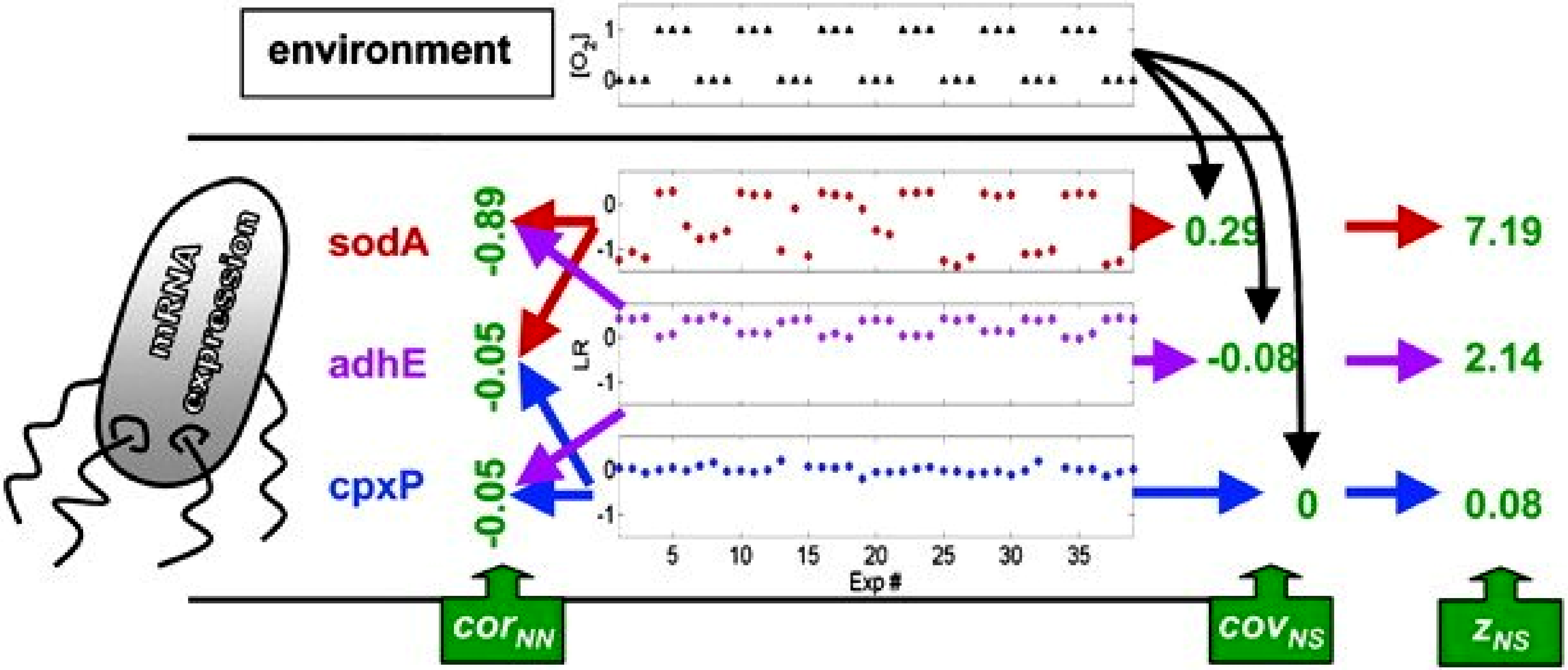
The origon network



Microarray data



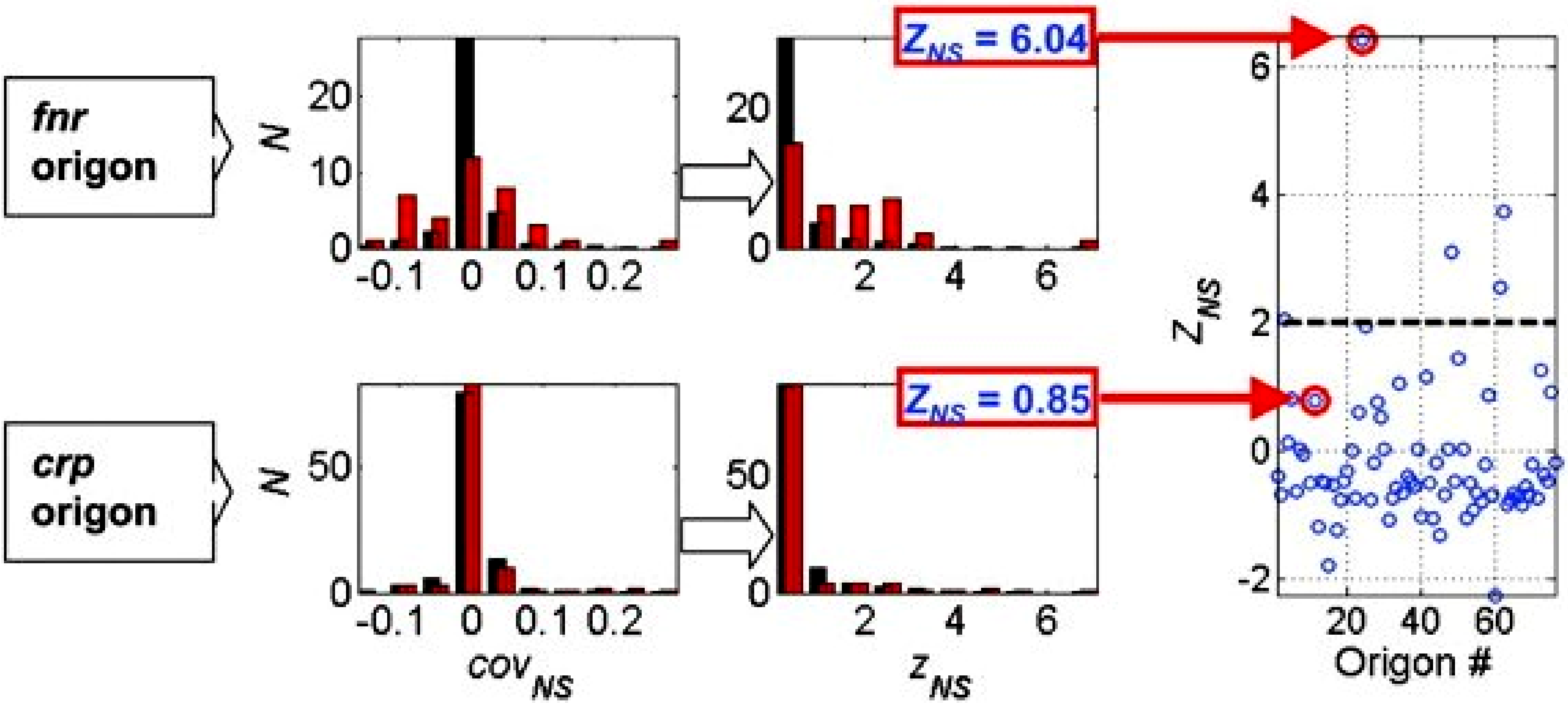
Response to extracellular oxygen



$$cor_{i,j} = \left\langle \frac{\sum_{c=1}^{N_c} LR_{r,c} LR_{r',c}}{\sigma_r \sigma_{r'}} \right\rangle_{r \in I, r' \in J}$$

$$\mathrm{cov}_i = \left\langle \sum_{c=1}^{N_c} LR_{r,c} S_c \right\rangle_{r \in I}$$

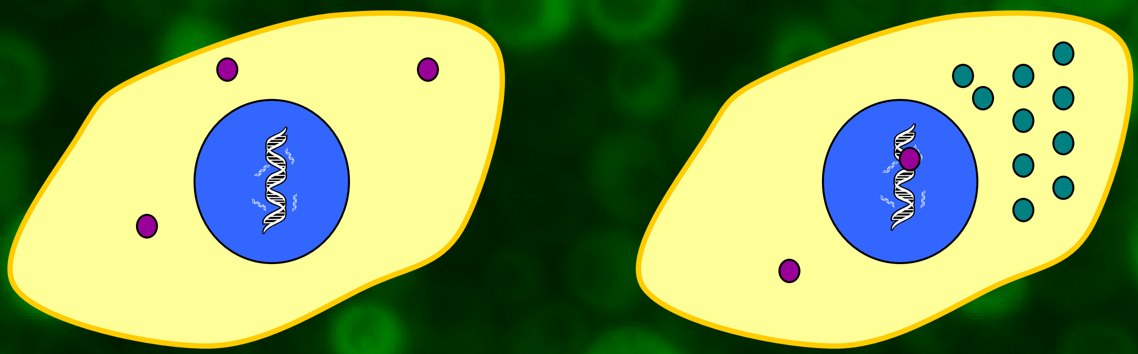
Significantly affected origons



$$z_{NS}(n)=\frac{\left|\mathrm{cov}_{NS}(n)-\mu_{NS}\right|}{\sigma_{NS}} \qquad \mu_O=\left\langle z_{NS}\right\rangle_O \qquad Z_{NS}=\frac{\mu_O-\mu_R}{\sigma_R}$$

Significantly affected origons: fnr, soxR, narL, rpoN, adiA-adiY

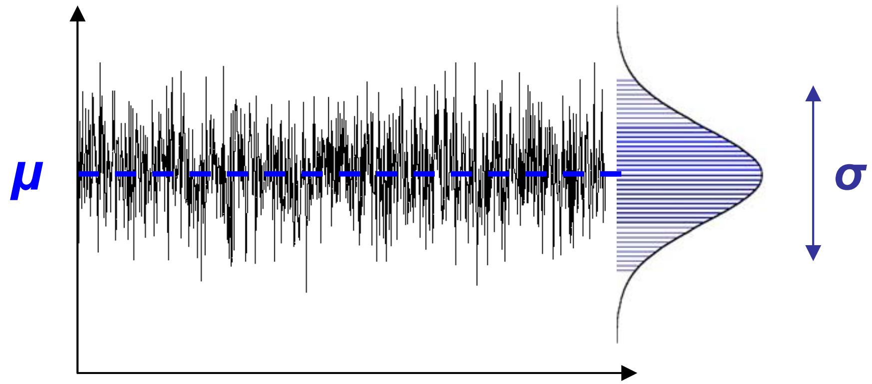
Gene Expression Noise (Part II)



Genetically identical yeast cells
Expressing a fluorescent reporter
Photo by Kevin F. Murphy (Boston U)

Measuring noise

Noise = CV = σ/μ



Noise =
Coeff. of variation =
CVR =
STD/mean

$$\frac{\sigma}{\mu}$$

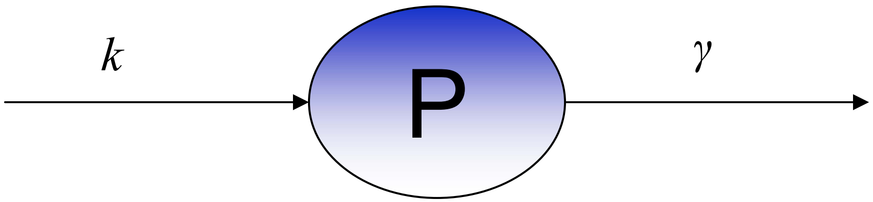
Elowitz et al., *Science* **297**, 1183 (2002)
Becskei et al., *Nat. Gen.* **37**, 937 (2005)
Colman-Lerner et al., *Nature* **437**, 699 (2005)

Noise strength =
Fano factor =
f =
Var/mean

$$\frac{\sigma^2}{\mu}$$

Ozbudak et al., *Nat. Gen.* **31**, 69 (2002)
Blake et al., *Nature* **422**, 633 (2003)
Raser & O'Shea, *Science* **304**, 1811 (2004)

A simple example: birth-death process



$\gamma = ln(2)/\tau$
τ: half-life

$$\frac{dP_N}{dt} = kP_{N-1} + \gamma(N+1)P_{N+1} - (k+\gamma N)P_N; \quad N \neq 0,\ P_{-1} = 0$$

Master equation

$$\frac{d\langle N\rangle}{dt} = k - \gamma\langle N\rangle \Rightarrow \mu(t) = \frac{k}{\gamma}\left(1 - e^{-\gamma t}\right)$$

Time evolution: mean

$$\frac{d\langle N^2\rangle}{dt} = -2\gamma\langle N^2\rangle + (\gamma + 2k)\langle N\rangle + k \Rightarrow \sigma^2(t) = \frac{k}{\gamma}\left(1 - e^{-\gamma t}\right)$$

Time evolution: variance

$$CV = \frac{\sigma(t)}{\mu(t)} = \frac{1}{\sqrt{\frac{k}{\gamma}\left(1 - e^{-\gamma t}\right)}}$$

CV = Coefficient of Variation

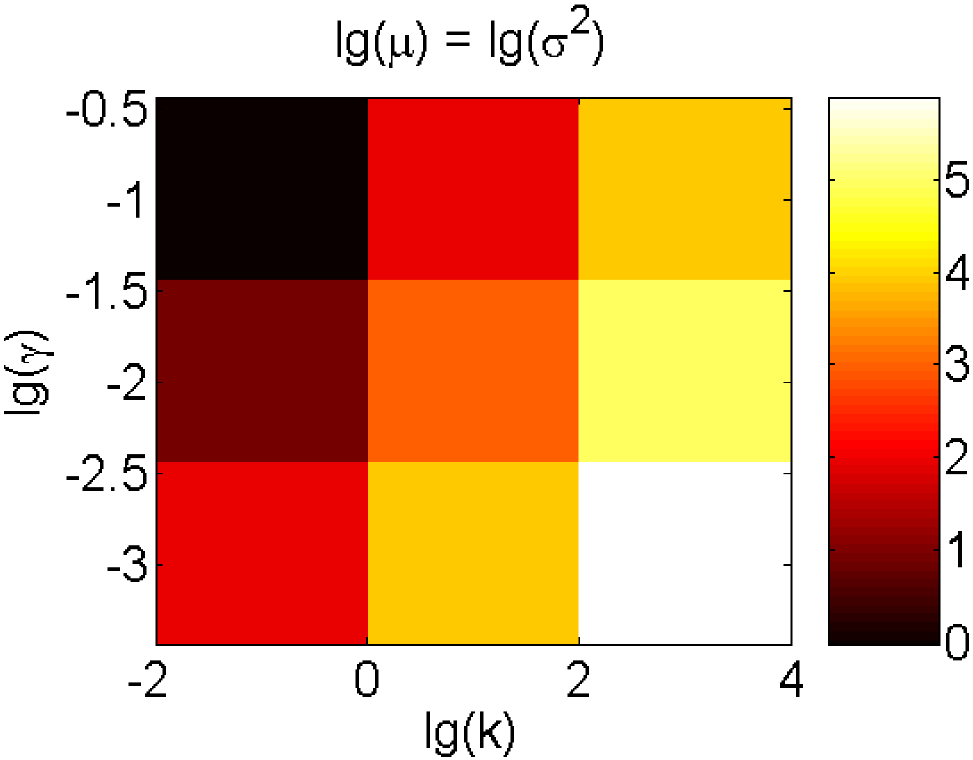
Time evolution: noise

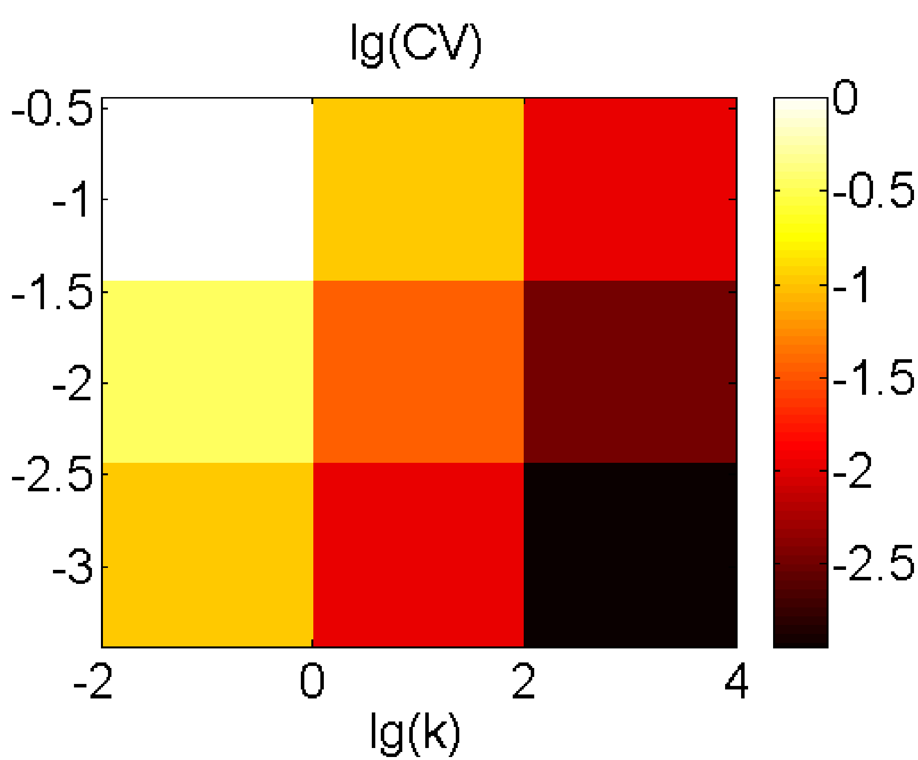
Stochastic Processes in Physics and Chemistry (North-Holland Personal Library) by N.G. Van Kampen

Steady-state: birth-death process

$$\mu_{SS} = \sigma_{SS}^2 = \frac{k}{\gamma}$$

$$CV_{SS} = \frac{\sigma_{SS}}{\mu_{SS}} = \sqrt{\frac{\gamma}{k}}$$





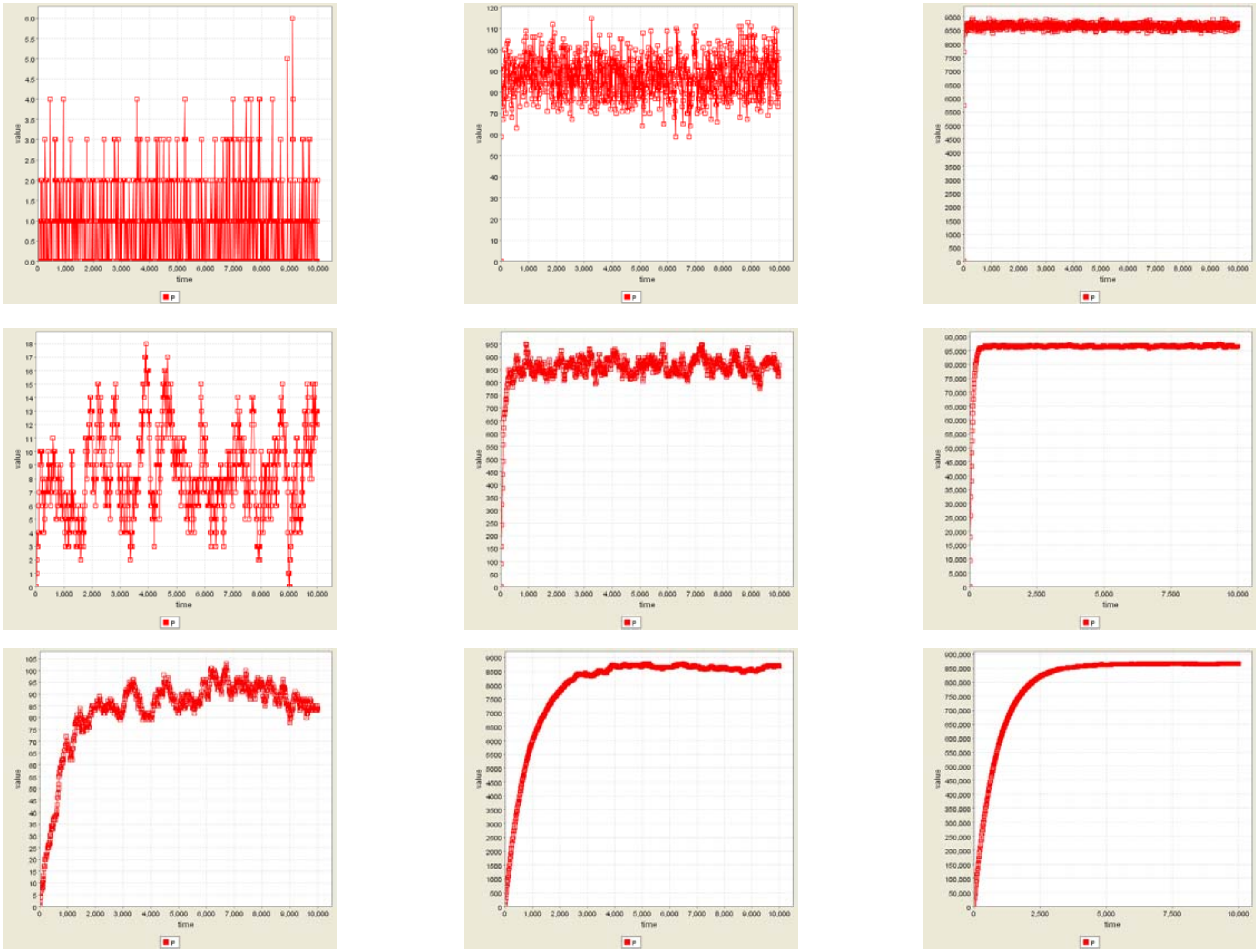
Stochastic simulation: birth-death process

k=0.1 *k=10* *k=1000*

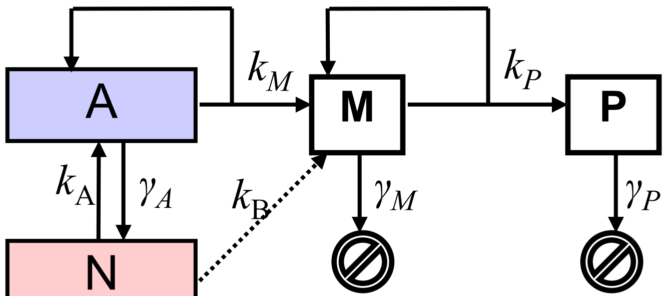
γ=ln(2)/6

γ=ln(2)/60

γ=ln(2)/600



Gene expression models



Kærn et al. *Nat. Rev. Genet.* **6**, 451-464 (2005)

Dynamics:

$$\dot{N} = -k_A N + \gamma_A A$$

$$\dot{A} = k_A N - \gamma_A A$$

$$\dot{M} = k_M A - \gamma_M M$$

$$\dot{P} = k_P M - \gamma_P P$$

Reactions:

$$N \xrightarrow{k_A} A$$

$$A \xrightarrow{\gamma_A} N$$

$$A \xrightarrow{k_M} A + M$$

$$M \xrightarrow{k_P} M + P$$

$$M \xrightarrow{\gamma_M} \varnothing$$

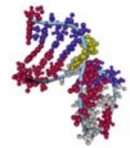
$$P \xrightarrow{\gamma_P} \varnothing$$

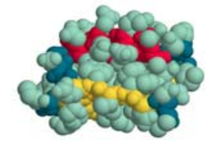
Steady state:

$$A_{SS} = \frac{k_A}{k_A + \gamma_A} \qquad M_{SS} = \frac{k_M}{\gamma_M} \frac{k_A}{k_A + \gamma_A}$$

$$N_{SS} = \frac{\gamma_A}{k_A + \gamma_A} \qquad P_{SS} = \frac{k_P}{\gamma_P} \frac{k_M}{\gamma_M} \frac{k_A}{k_A + \gamma_A}$$

Sources of noise





Process #1: Promoter binding (DNA) → Process #2: Transcription (mRNA) → Process #3: Translation (Protein)

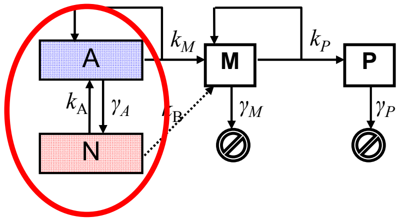
$$\frac{\sigma_1^2}{\langle n_1 \rangle^2} = \frac{\lambda_1^-}{\lambda_1^+} = \frac{\gamma_A}{k_A}$$

$$\frac{\sigma_2^2}{\langle n_2 \rangle^2} = \frac{1}{\langle n_2 \rangle} + \frac{\lambda_1^-}{\lambda_1^+}\frac{\tau_1}{\tau_2 + \tau_1}$$

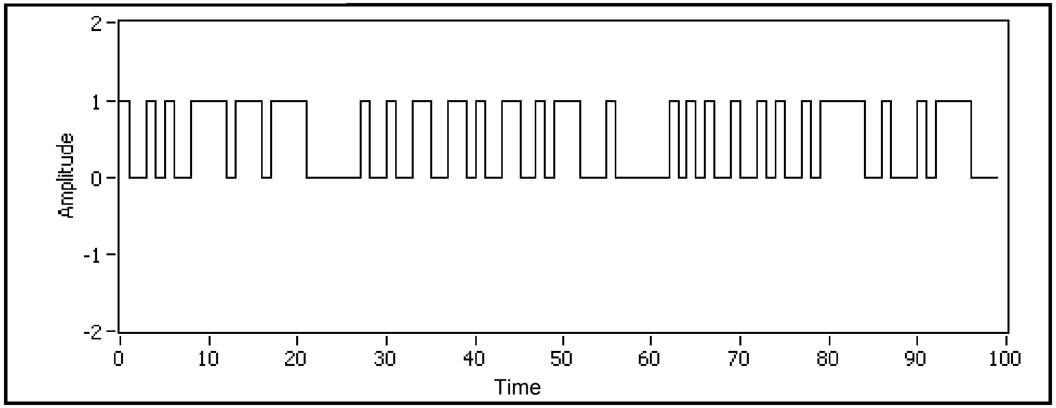
$$\overbrace{\frac{\sigma_3^2}{\langle n_3 \rangle^2}}^{\text{Total protein noise}} = \overbrace{\underbrace{\frac{1}{\langle n_3 \rangle}}_{\text{Poisson}}}^{\text{From individual births and deaths}} + \overbrace{\underbrace{\frac{1}{\langle n_2 \rangle}}_{\text{Poisson}} \underbrace{\frac{\tau_2}{\tau_3 + \tau_2}}_{\text{One-step time-averaging}}}^{\text{From spontaneous mRNA noise}} + \overbrace{\underbrace{\frac{1 - P_{\text{on}}}{\langle n_1 \rangle}}_{\text{Binomal}} \underbrace{\frac{\tau_2}{\tau_2 + \tau_3}\frac{\tau_1}{\tau_1 + \tau_3}\frac{\tau_1 + \tau_3 + \tau_1\tau_3/\tau_2}{\tau_1 + \tau_2}}_{\text{Two-step time-averaging}}}^{\text{From forced mRNA noise, originating in gene activation-inactivation}}$$

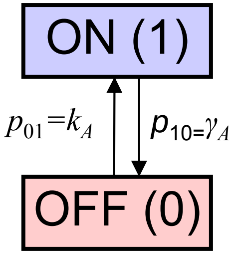
Paulsson, J., *Nature* **427**, 415-8 (2004).

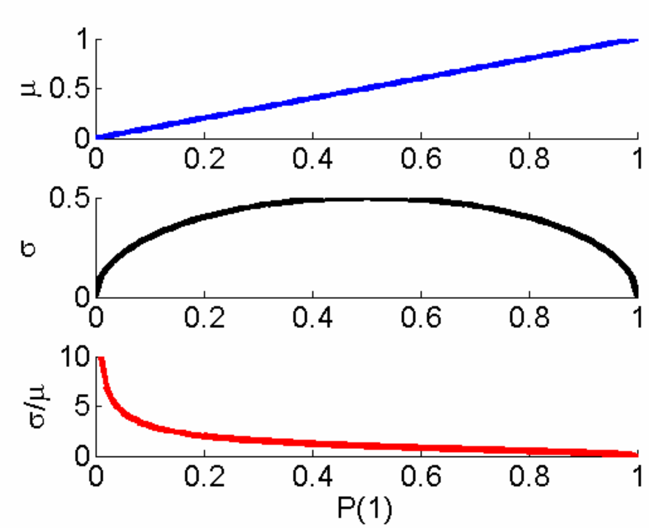
Noise at the promoter: Random Telegraph Process



A+*N*=const=1





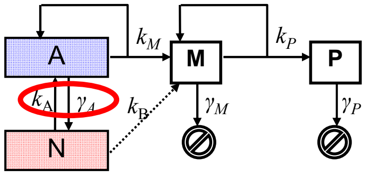


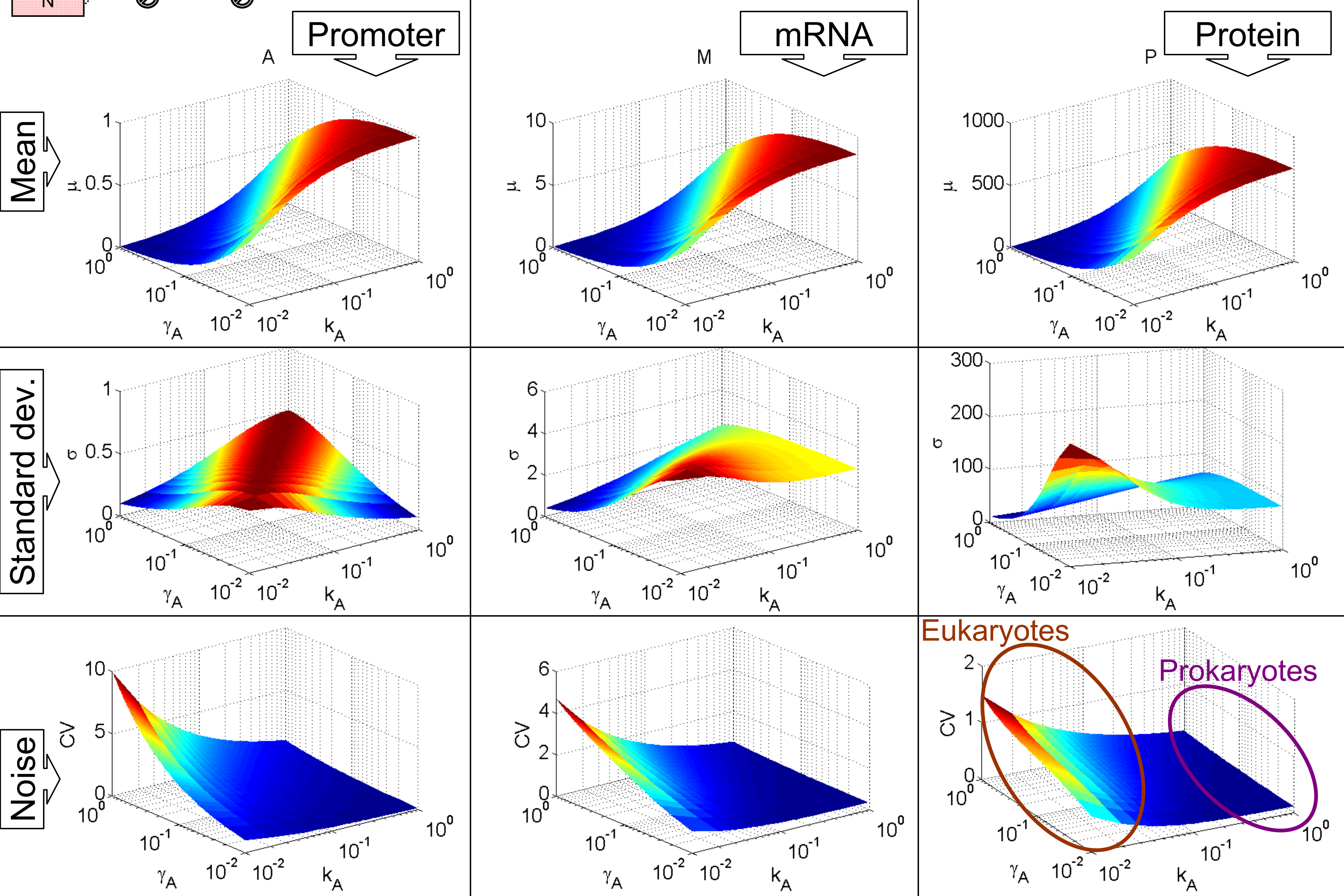
$$\mu = \frac{p_{01}}{p_{01} + p_{10}}$$

$$\sigma = \frac{\sqrt{p_{01} p_{10}}}{p_{01} + p_{10}}$$

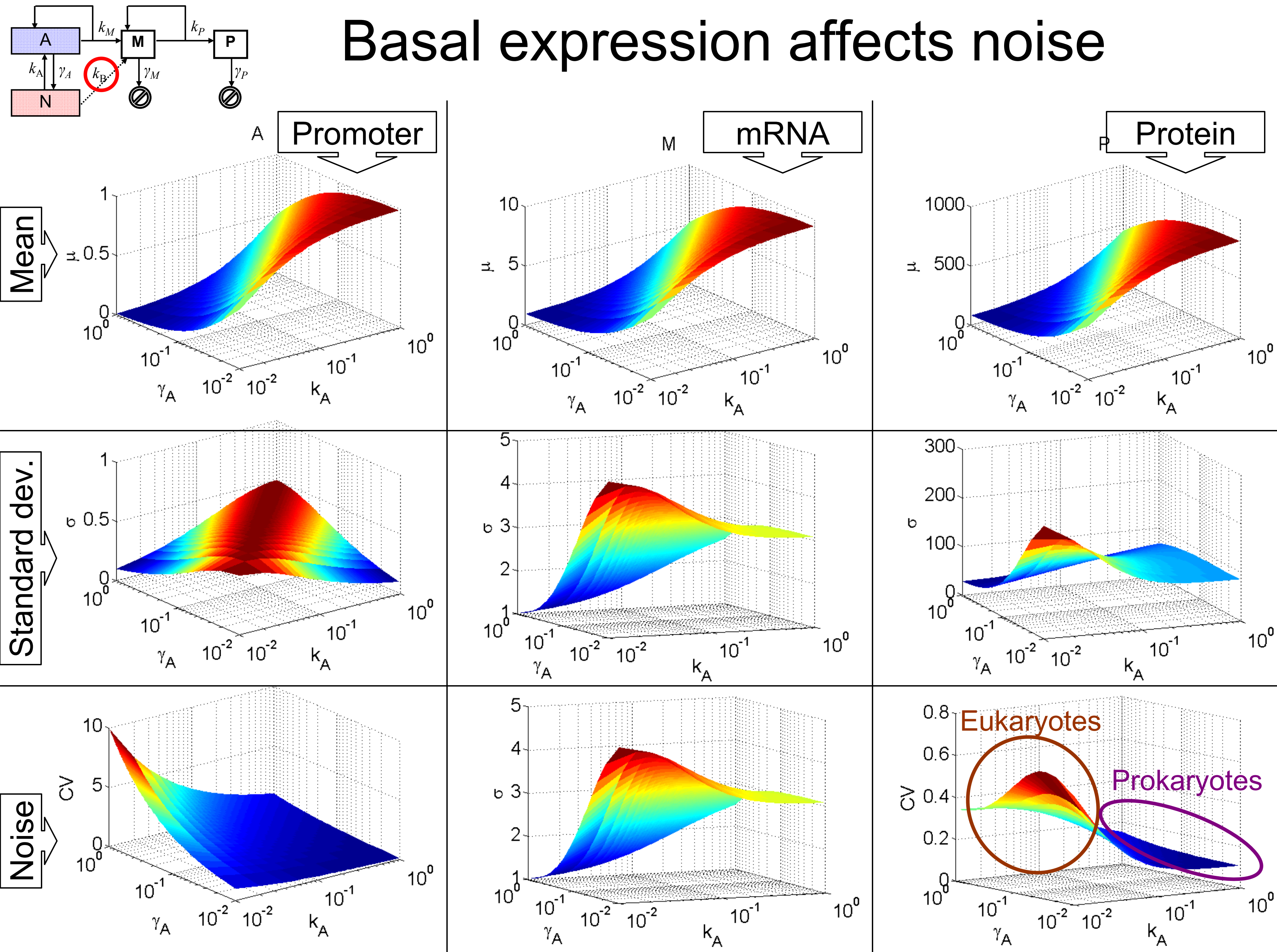
$$\frac{\sigma}{\mu} = \sqrt{\frac{p_{10}}{p_{01}}}$$

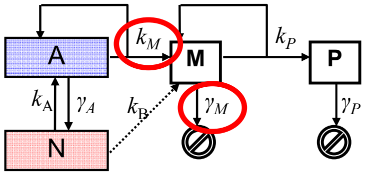
Promoter dynamics affects noise



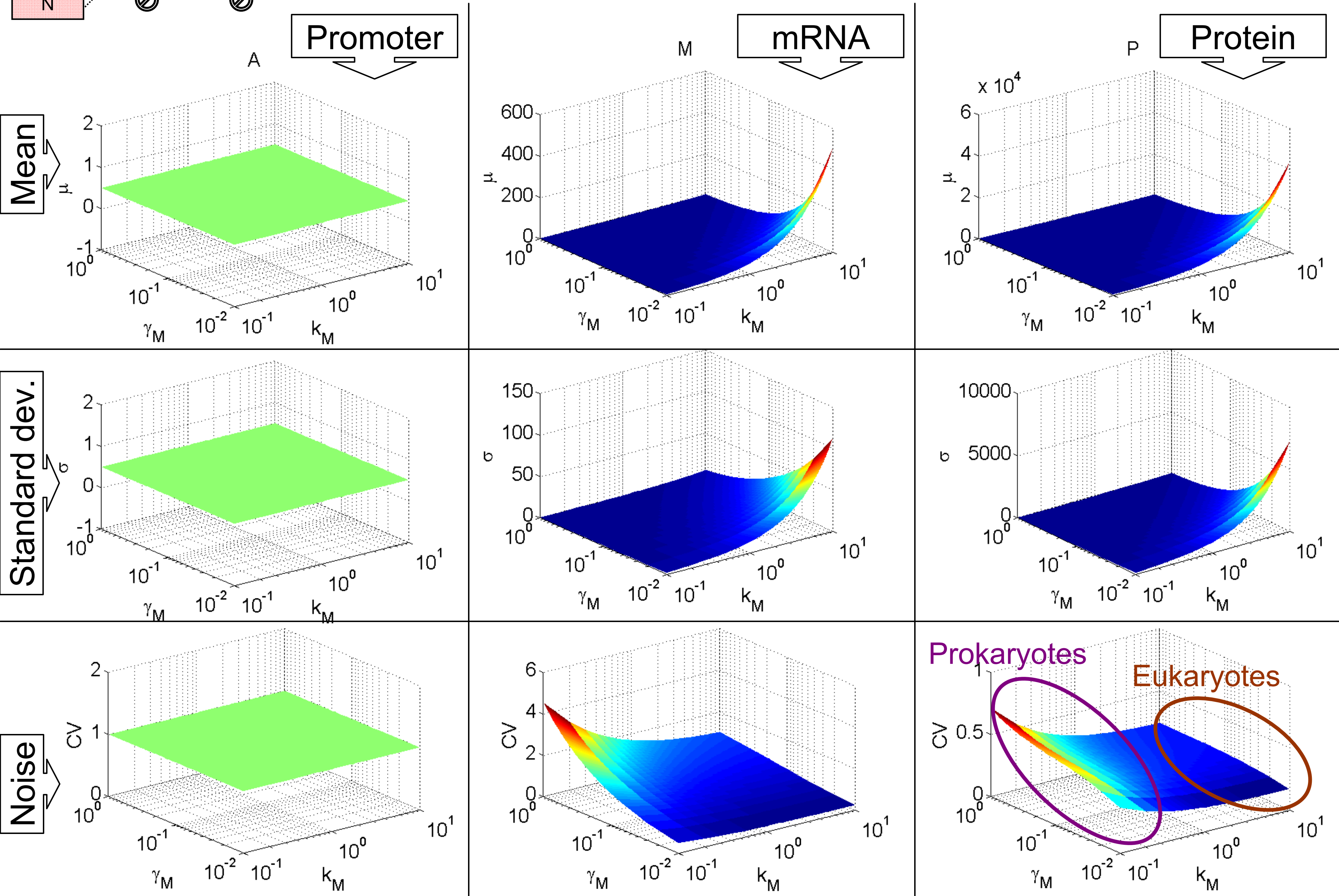


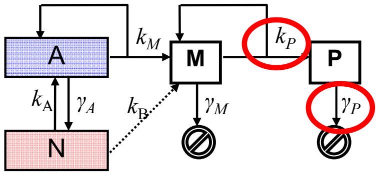
Basal expression affects noise



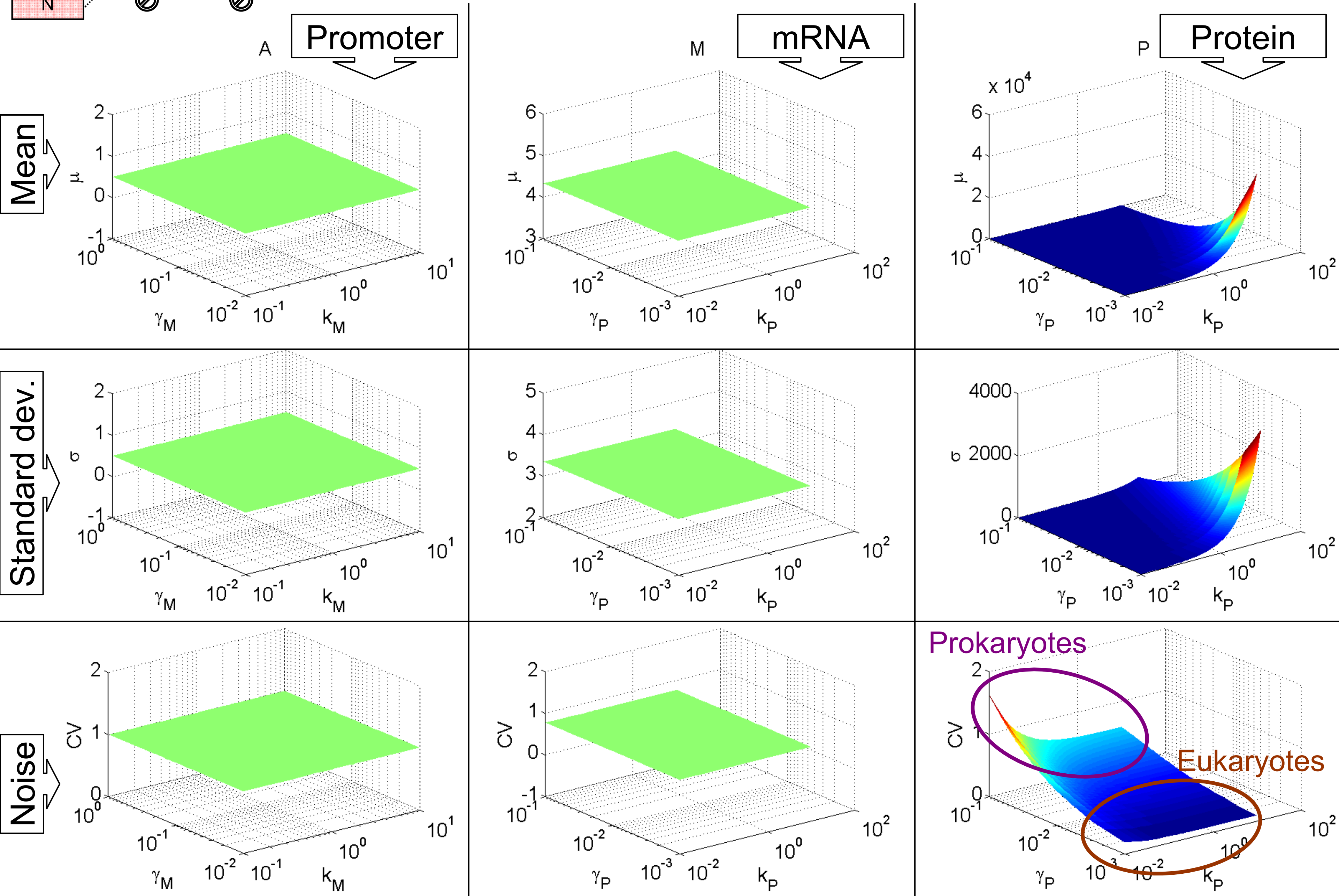


mRNA birth/death affects noise

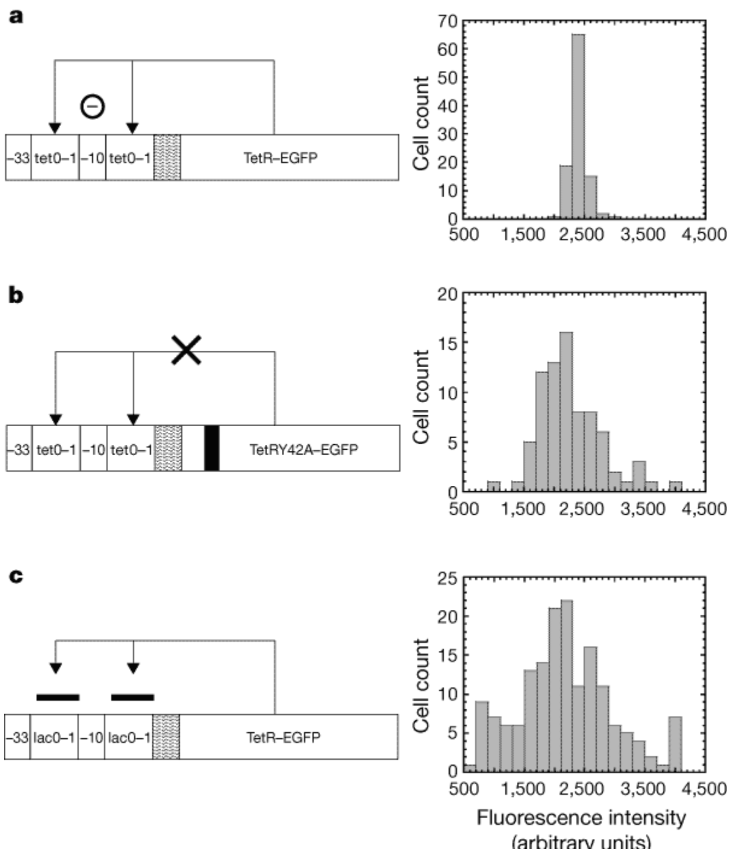




Protein birth/death affects noise

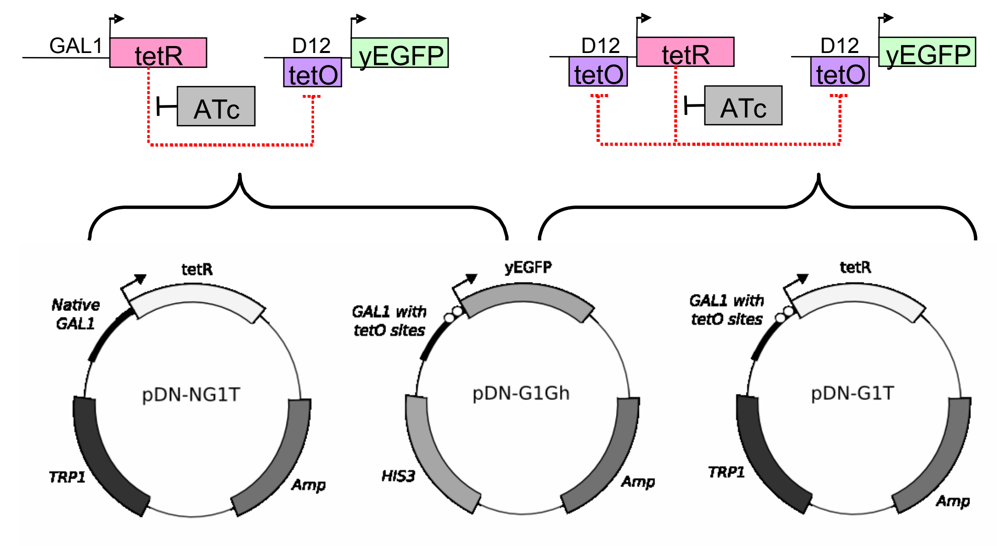


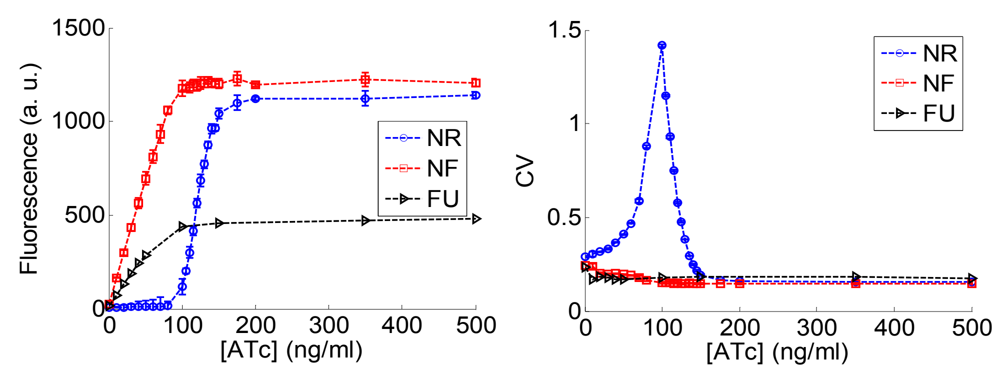
Noise and negative feedback



Becskei et al., *Nature* **405**, 590 (2000).

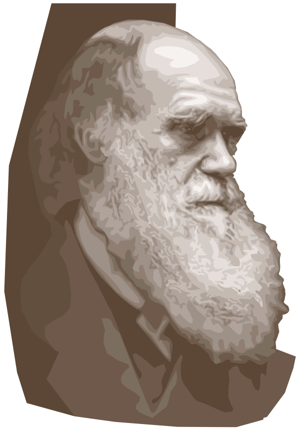
Negative Regulation (NR) Negative Feedback (NF)





Survival and evolution in a changing environment (Part III)

Evolutionary theory



- Evolution is the process through which the heritable traits of a biological population change from one generation to the next.

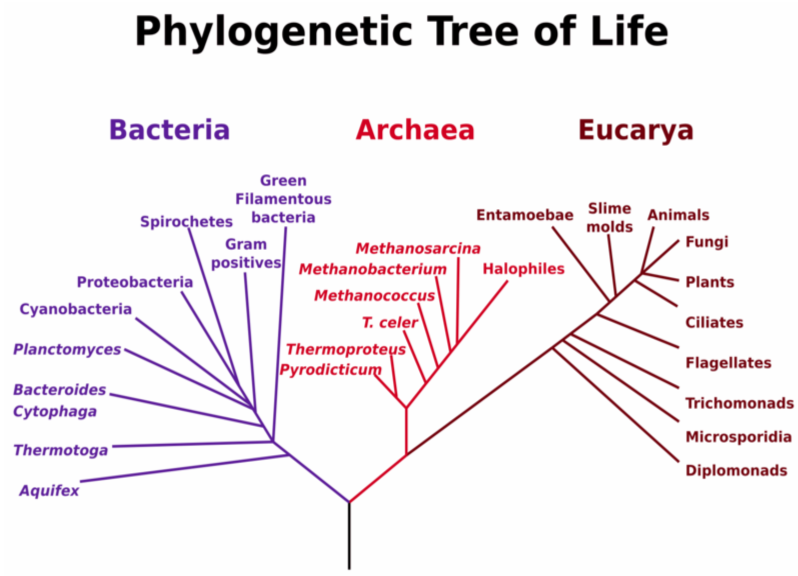
-Charles Darwin: The Origin of Species (1859)

-Gregor Mendel: Experiments on Plant Hybridization (1866)

-Population Genetics (1900-1930)

-Modern Synthesis: Fisher, Haldane, Wright, Huxley, Mayr (1930-1950)

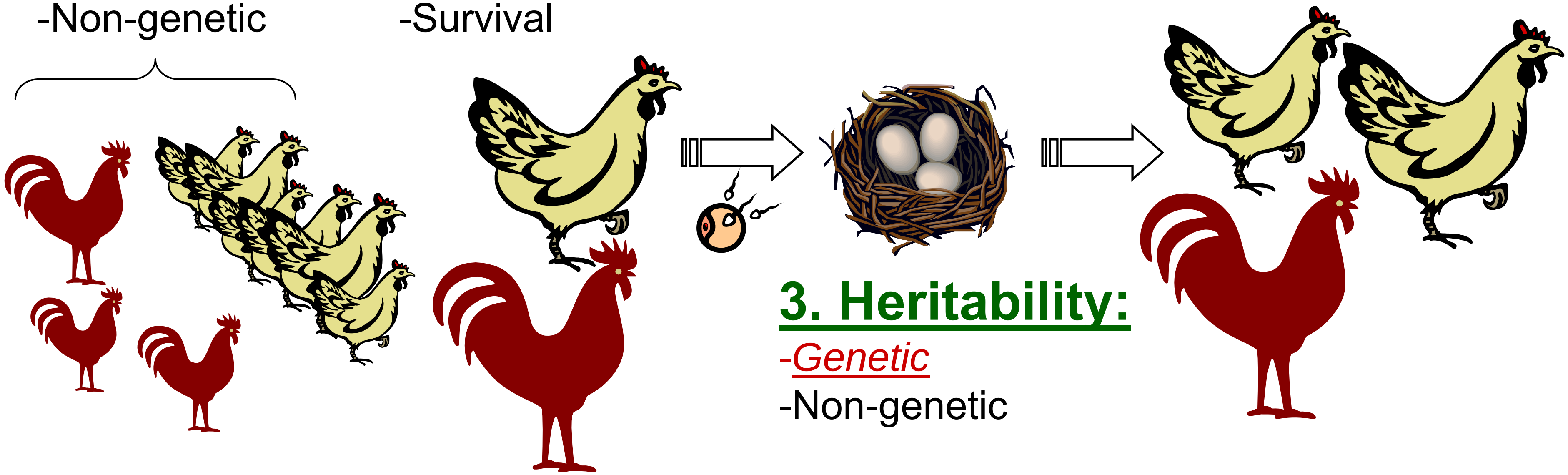
-Neutral Theory: M. Kimura (1960-1980)



Requirements for evolution under selection

1. Variation:
-Genetic
-Non-genetic

2. Selection:
-Fitness + Competition
-Survival



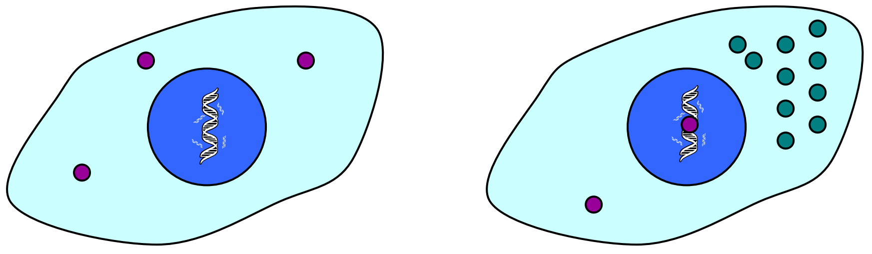
3. Heritability:
-Genetic
-Non-genetic

This is the **textbook-understanding**, focusing on **multicellular species**.

- **Phenotype:**
  - Observed quality (size, color, shape, ...)
  - Subject to selection
- **Genotype:**
  - DNA sequence
  - Not directly selected
  - Makes phenotypes heritable

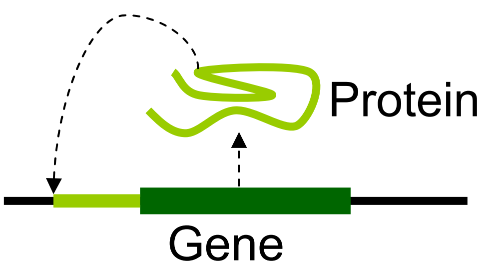
Sources of cell-cell variation

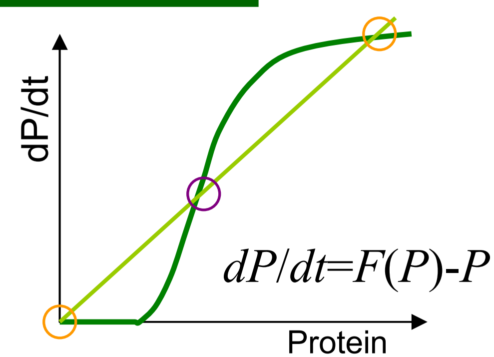
- Variation is required for evolution
- The source of phenotypic variation can be:
  - Genetic
    - Gene mutations (amplifications,deletions, insertions,…)
  - Epigenetic
    - DNA methylation
    - Chromatin modification
  - Non-genetic (**NOISE**)
    - Low intracellular concentrations

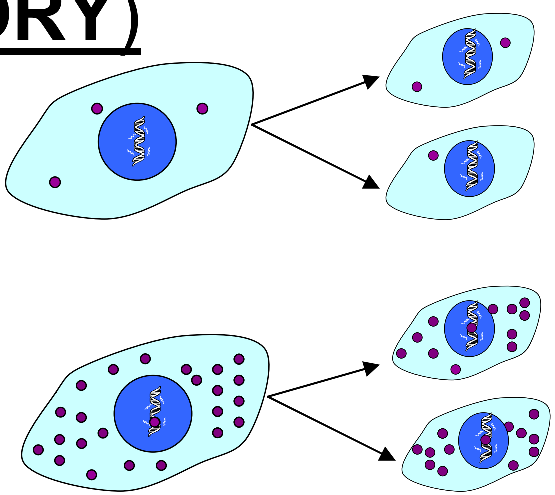


Heritability of cellular phenotypes

- Heritability is required for evolution
- The source of heritability can be:
  - <u>Genetic</u>
    - Gene mutations (amplifications,deletions, insertions,…)
  - <u>Epigenetic</u>
    - DNA methylation
    - Chromatin modification
  - <u>Non-genetic (**CELLULAR MEMORY**)</u>
    - Positive feedback

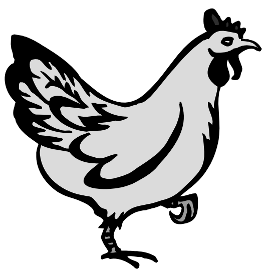


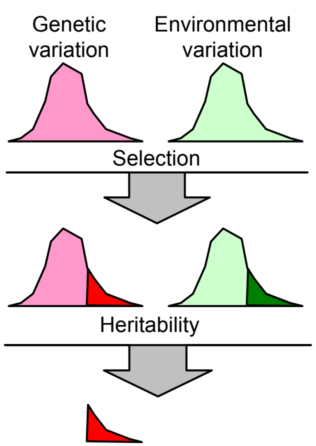




Evolution at the single-cell level

Multicellular organisms

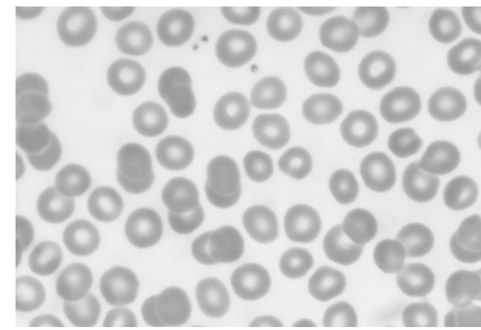


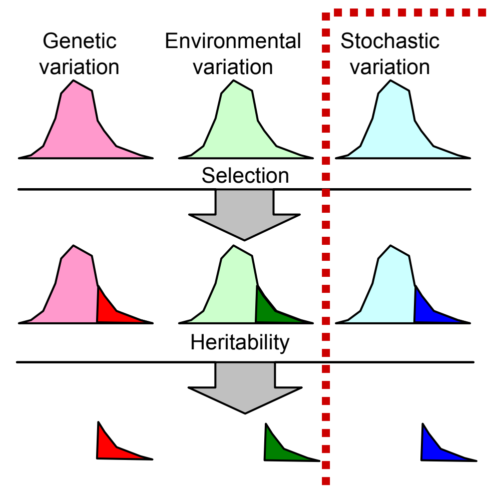


Differences

Generation time:	
Years	Minutes
Eff. population size:	
~100	$>10^6$
Reproduction:	
Sexual	Asexual
Heritable variation	
Only genetic	Genetic+ non-genetic

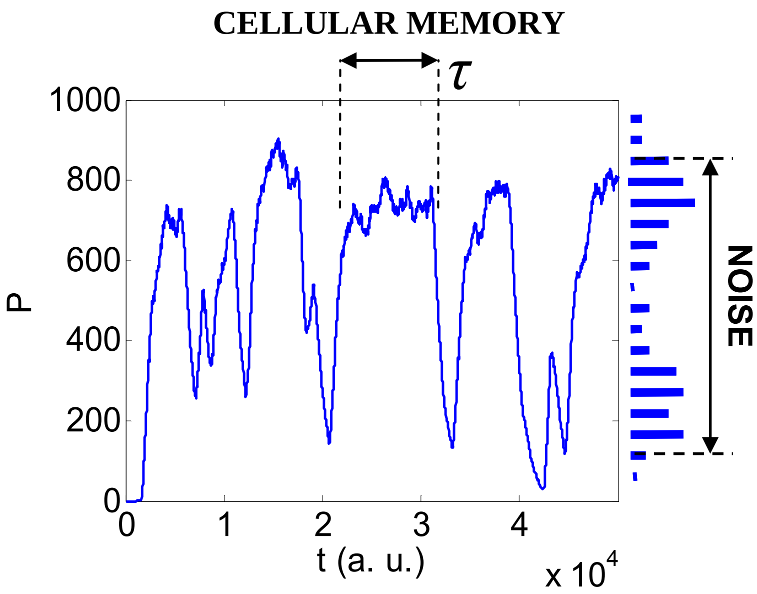
Single cells

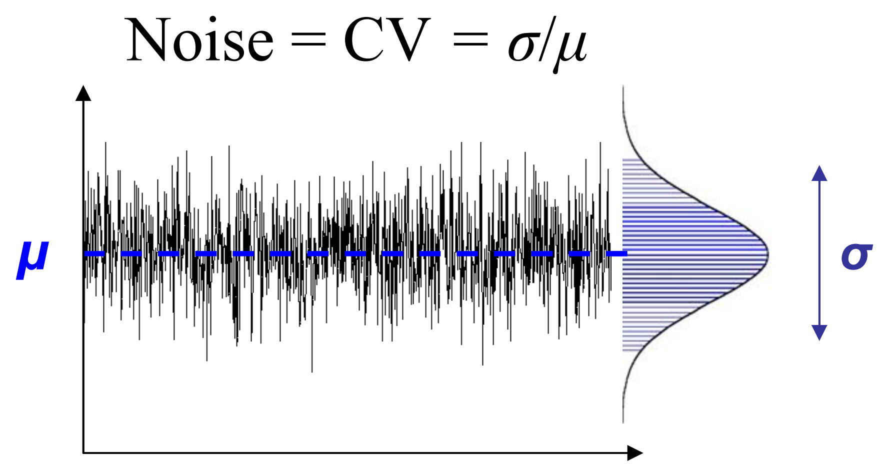




-Genome-scale data available (yeast & *E. coli*)
-Easy to manipulate genetically

CELLULAR MEMORY and NOISE





Noise:

-Quantifies non-genetic deviations from the population mean

-Measured by the Coefficient of Variation (CV, standard deviation / mean)

Cellular memory:

-The capacity of cell lineages to maintain deviant states non-genetically over time.

-It is the inverse of the rate of switching between phenotypic states ($1/\tau$)

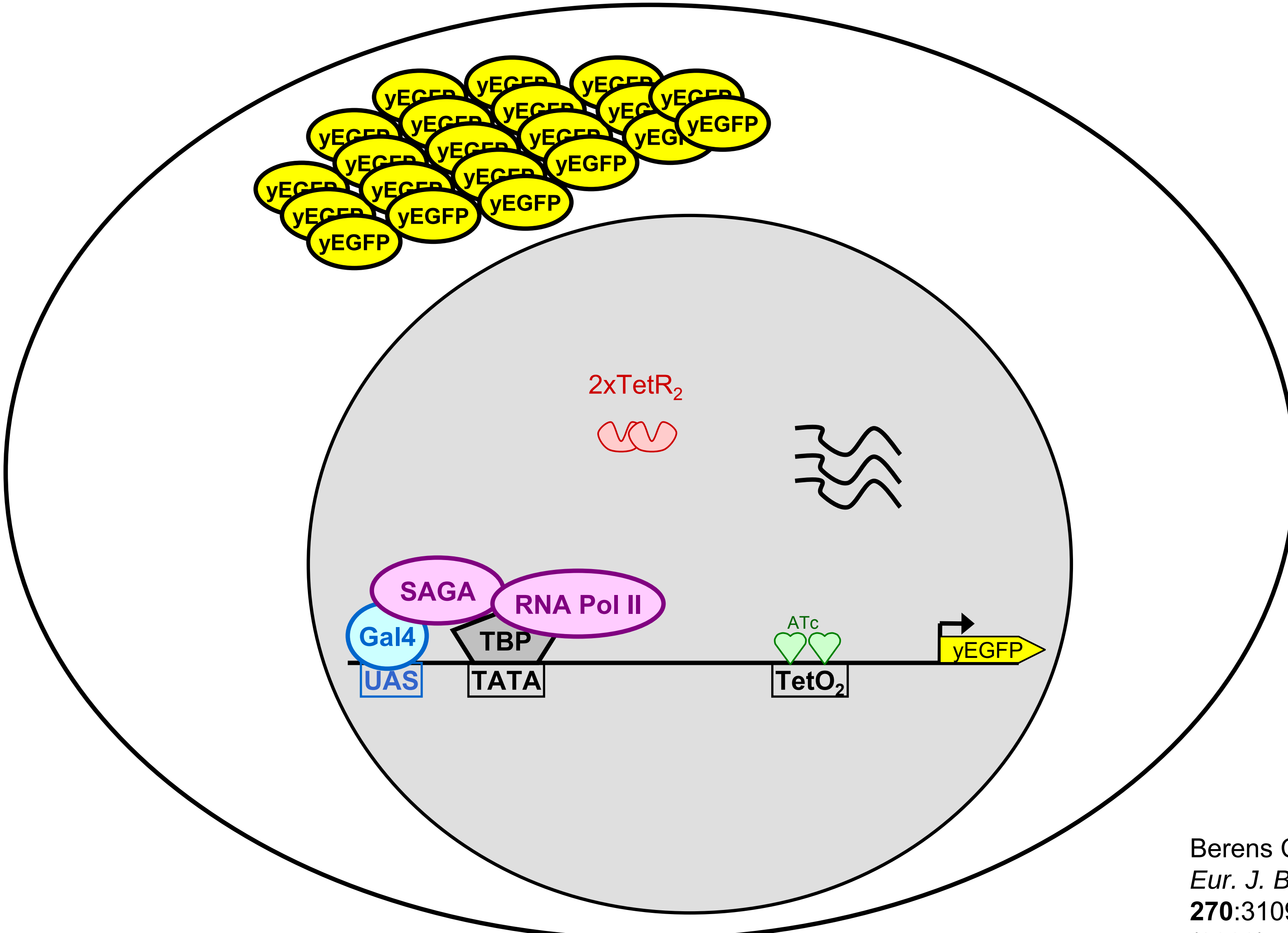
How do they affect survival and evolution at the single cell level?

Noise and phenotype

- To study the phenotypic effect of noise, we need to control it experimentally
- Controlling noise also affects other cellular properties (e.g., gene expression mean)
- Therefore, we need to uncouple the control of noise from the control of mean

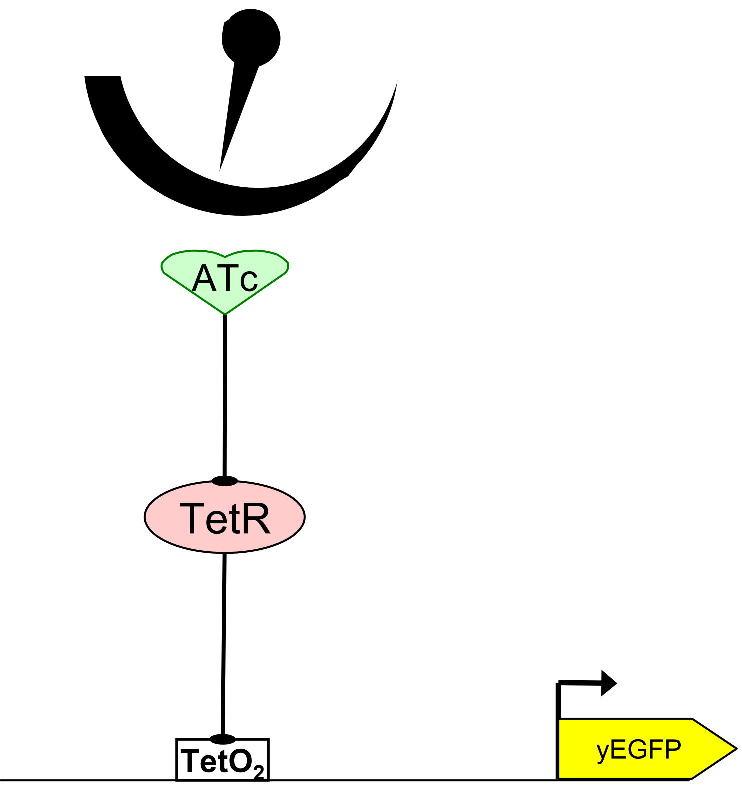
Photo by Kevin F. Murphy

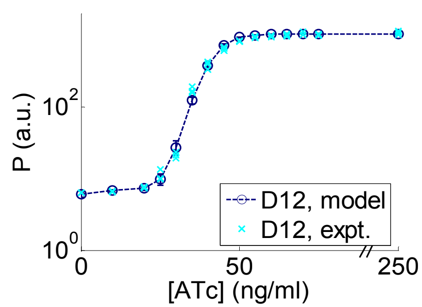
The TetR-repressible GAL1 promoter

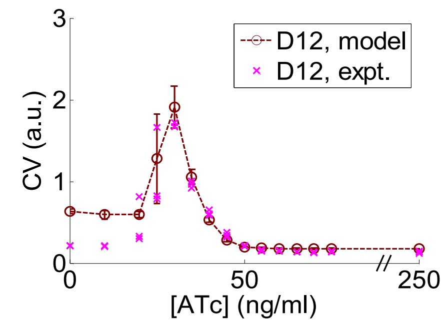


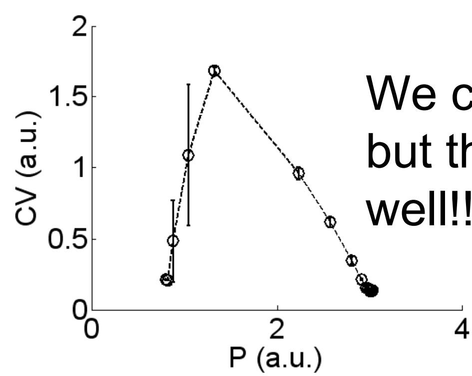
Berens C, Hillen W, *Eur. J. Biochem.* **270**:3109-21 (2003). Review.

Changing noise affects the mean





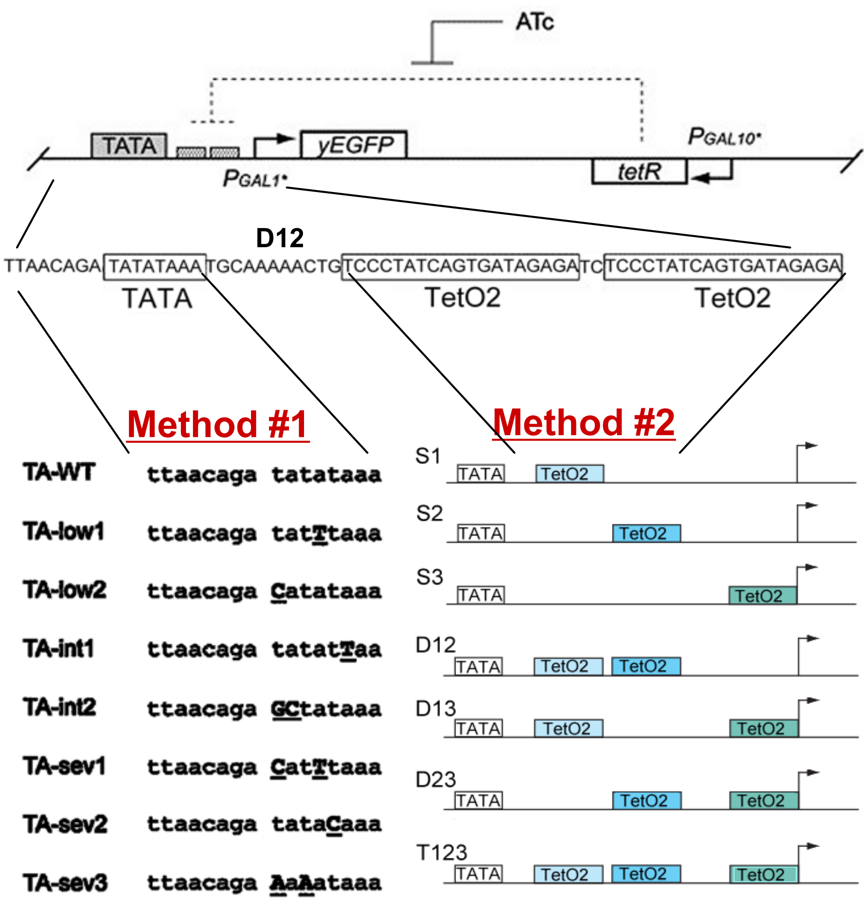




We can change the noise…
but the mean changes as
well!!!

Blake WJ, Balázsi G, Kohanski MA, Isaacs FJ, Murphy KF, Kuang Y, Cantor CR, Walt DR, Collins JJ, *Mol Cell* **24**, 853-865 (2006).
Murphy KF, Balázsi G, Collins JJ, *PNAS* **104**, 12726-31 (2007)

Uncoupling the noise and the mean

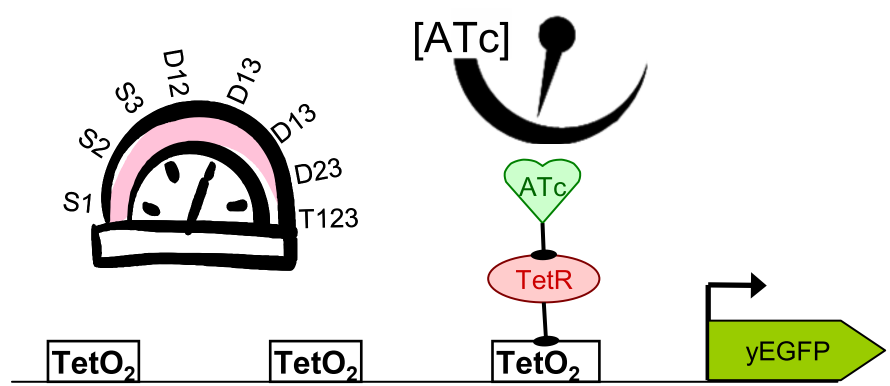


-The goal of these methods is to establish different dependence of the noise on the mean

-Method #1 consists of mutating the TATA box, decreasing the rate of promoter inactivation

-Method #2 consists of changing the number of repressor binding sites

-Mathematical models accompany both methods



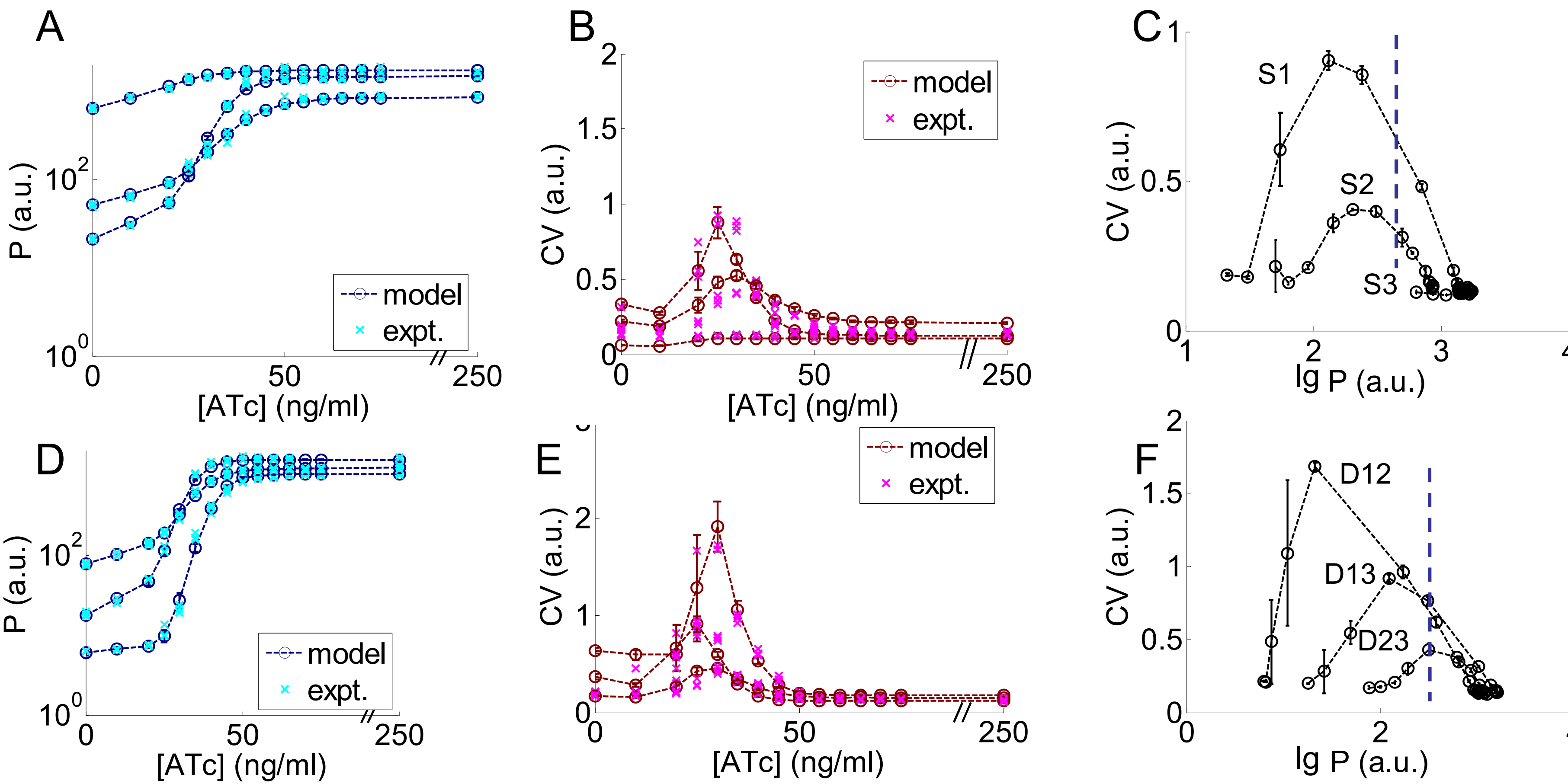
See also:

-Smith, Sumner & Avery, *Mol. Microb*. 2007

-Maamar, Raj & Dubnau, *Science* 2007

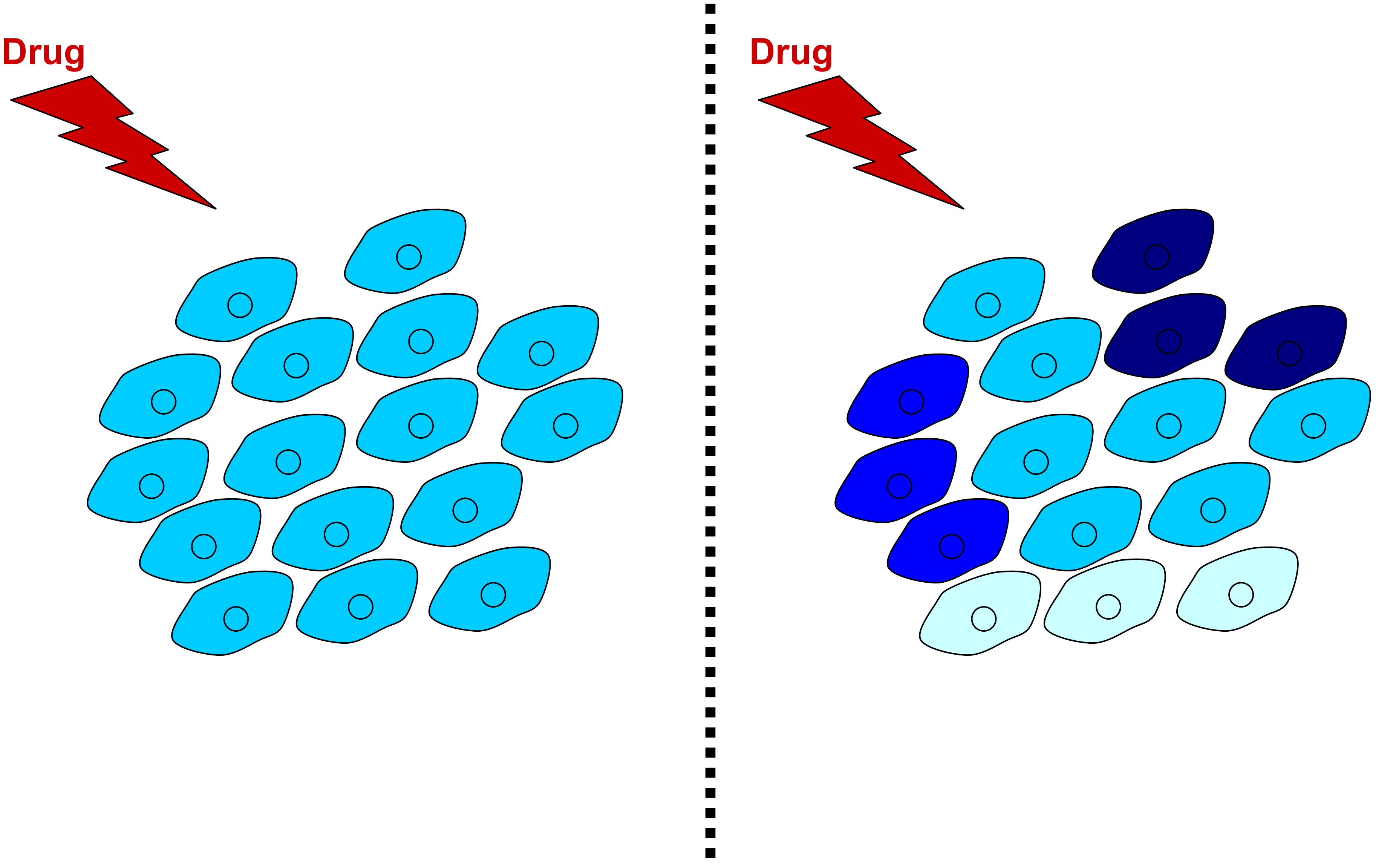
-Süel et al., *Science* 2007

Independent control of the noise and the mean

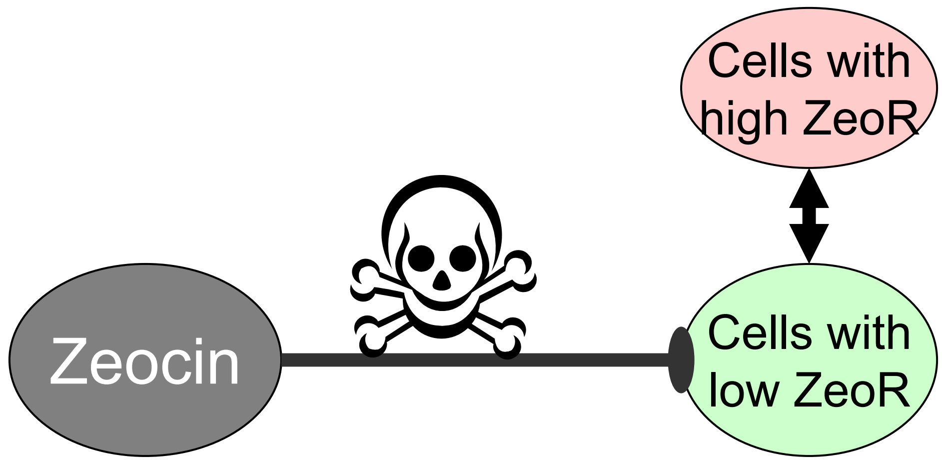


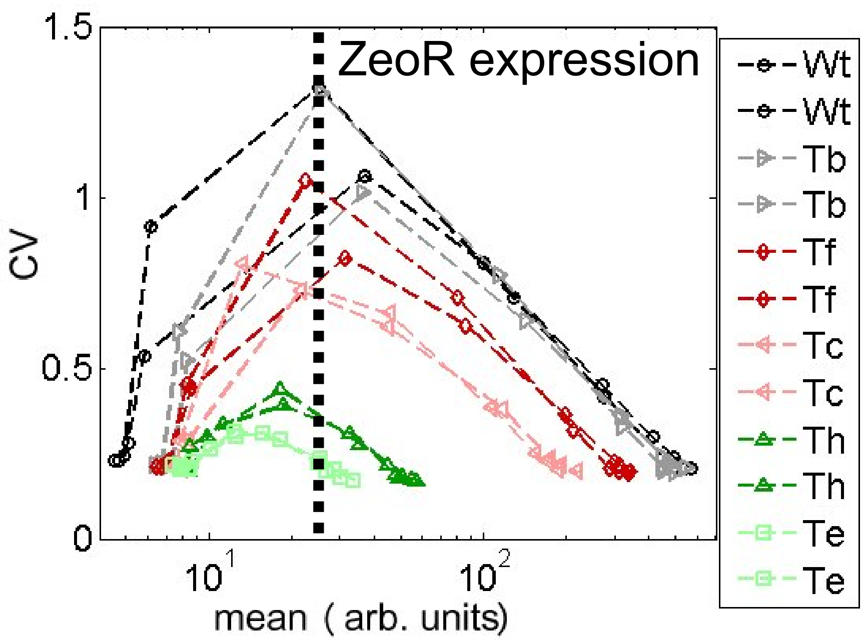
Blake WJ, Balázsi G, Kohanski MA, Isaacs FJ, Murphy KF, Kuang Y, Cantor CR, Walt DR, Collins JJ, *Mol Cell* **24**, 853-865 (2006).
Murphy KF, Balázsi G, Collins JJ, *PNAS* **104**, 12726-31 (2007)

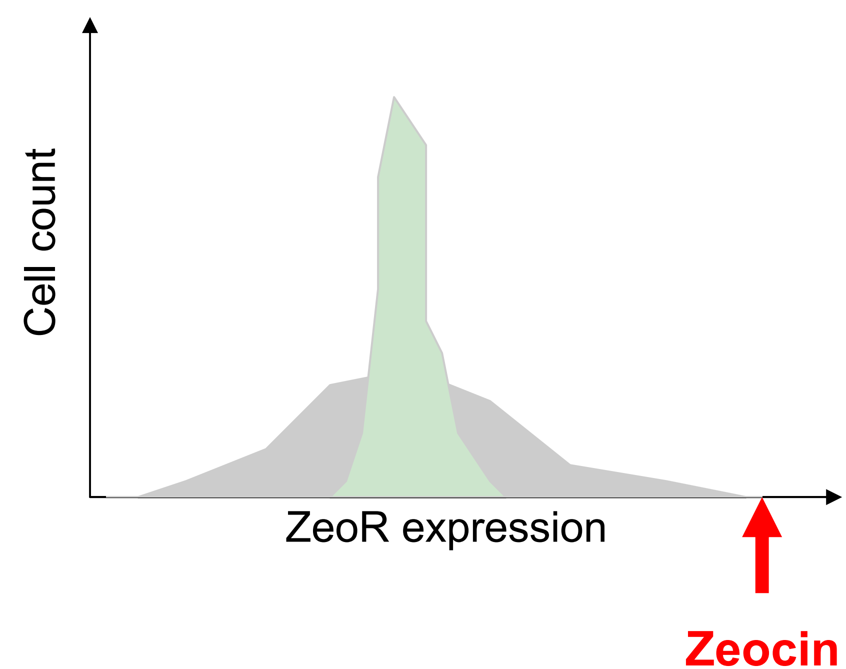
Noise and drug resistance



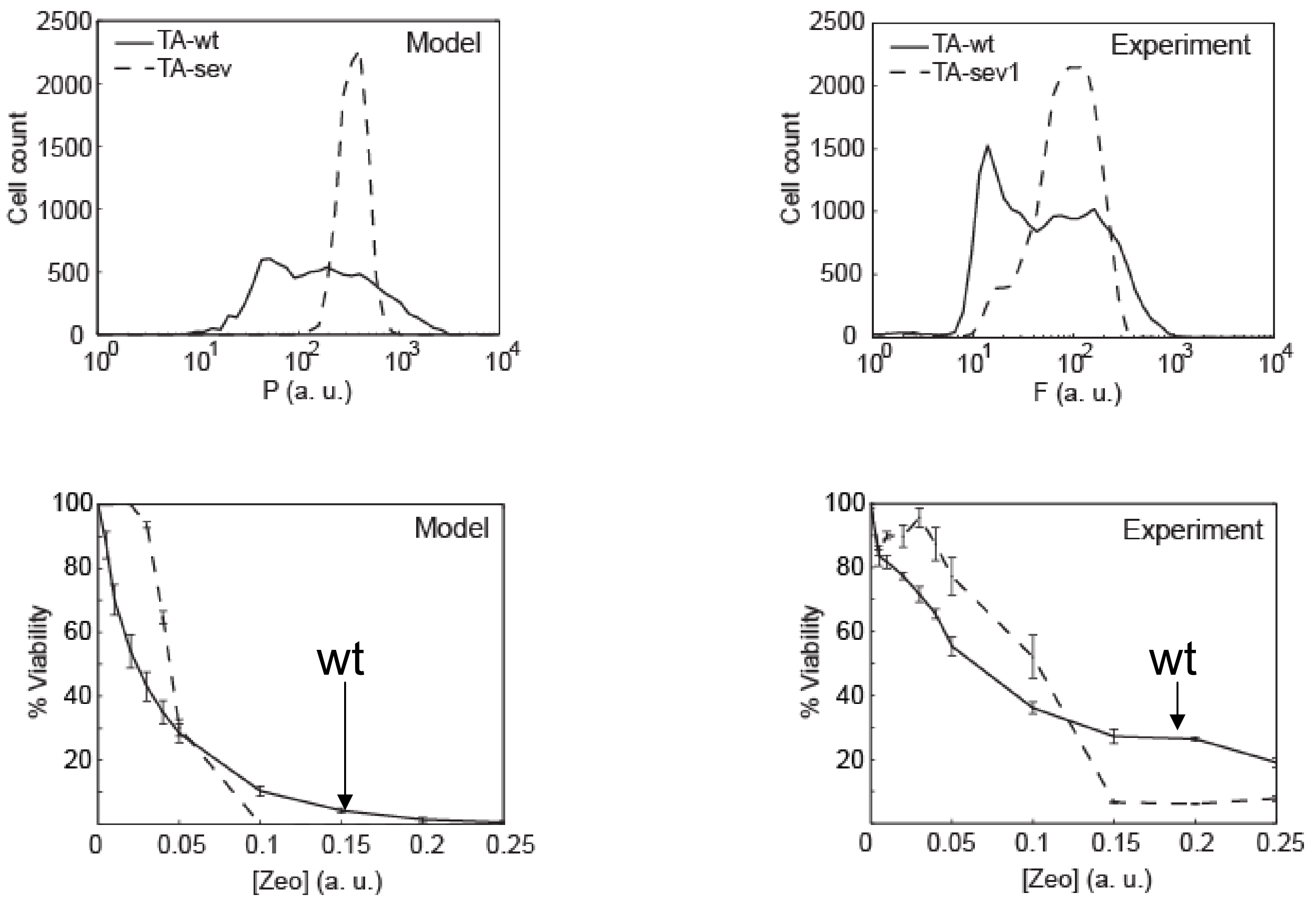
Controlling drug resistance noise







Noise aids survival during drug treatment



Blake WJ, Balázsi G, Kohanski MA, Isaacs FJ, Murphy KF, Kuang Y, Cantor CR, Walt DR, Collins JJ, *Mol Cell* **24**, 853-865 (2006)

CHIP Lab


Bhaskar Dutta
Dmitry Nevozhay
Rhys Adams
Vinícius Bonato
Allison Heath

Thank you!

- CHIP lab
- Gordon B. Mills
- Jordan Gutterman

THE UNIVERSITY OF TEXAS MD ANDERSON CANCER CENTER *Making Cancer History®*

- ITES Networks Cluster
- Tim F. Cooper
- Kevin Bassler
- Bogdan Dănilă

UNIVERSITY of HOUSTON

- Allison Heath
- Lydia Kavraki

RICE

- William Blake
- Kevin Murphy
- Jim Collins
- NIH, NSF

Applied Biodynamics Laboratory

BOSTON UNIVERSITY

CBD Center for BioDynamics Boston University

- Zoltán Oltvai

University of Pittsburgh School of Medicine Department of Pathology

University of Pittsburgh

- Albert-László Barabási

Northeastern University