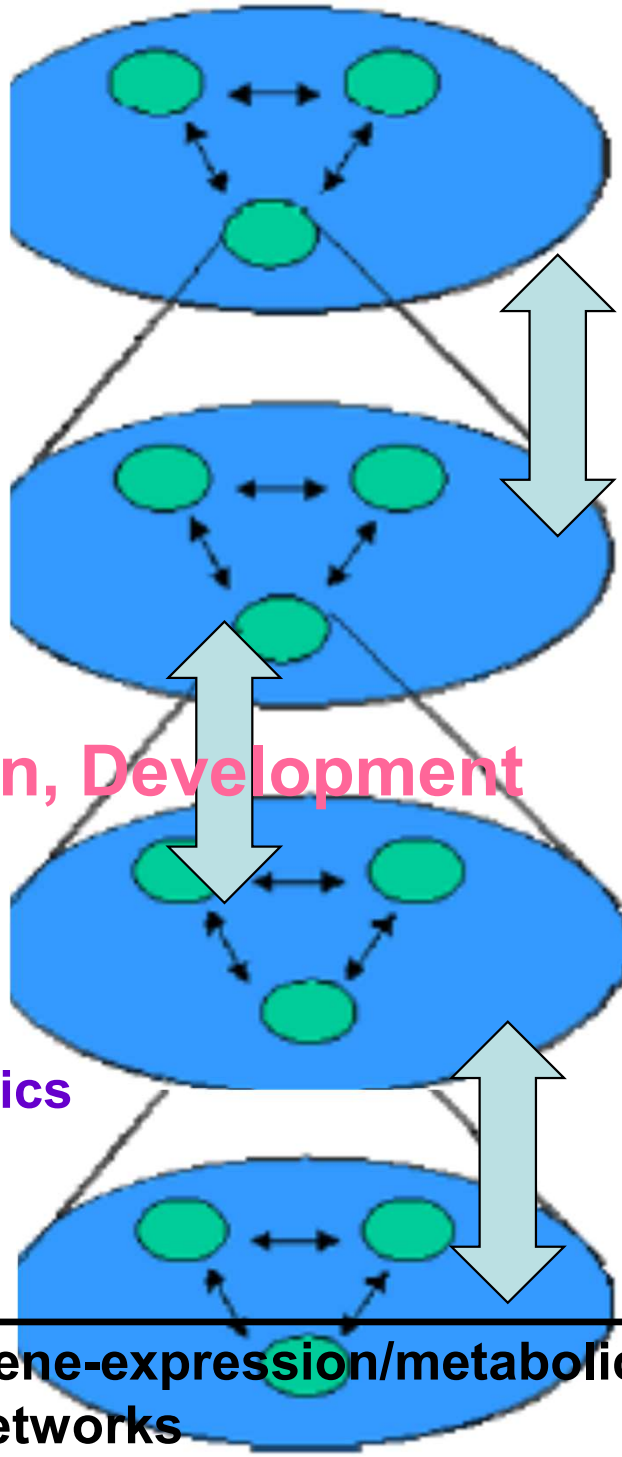
Consistency between hierarchical levels (+collapse)



Ecosystem

Multicellular organism

Cell-Tissue

Differentiation, Development

Cell

Cell-growth vs Intracellular dynamics

Adaptation→

Molecule

Gene-expression/metabolic networks

Extract general basic properties of life, based on this consistency principle

Diversity, Symbiosis

Evolution

Genotype ⟷ Phenotype

Molecule replication vs cell reproducti

Reproduction

Consistency with such distinct spatiotemporal scales?

- Grand Challenge:

Cell --- very high-dimensional dynamical systems (~5000 proteins for bacteria etc.)

- Can we understand it?
- Recall thermodynamics : huge-dimensional molecular dynamics, but described by few degrees ← restricting to *equilibrium*
- From high-dimensional dynamics of cell, surprisingly low-dimensional structure is extracted, with deep linearity ← restricting to *steady-growth states*: Valid after evolution, not any high-dim dynamical systems

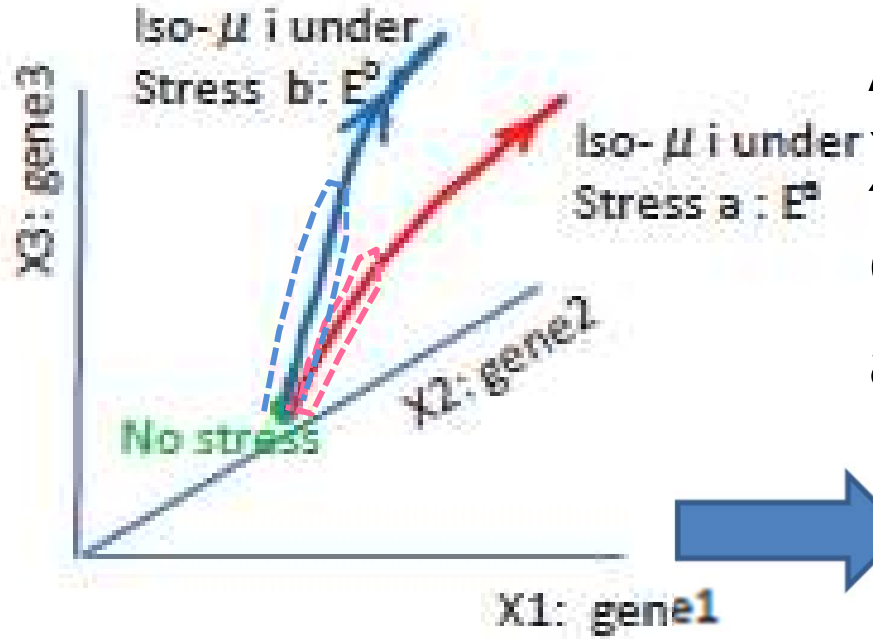
steady-growth →universal constraint

many components (few thousand proteins,,) in a cell steady -> all the components have to be roughly doubled before division)

Ni(i=1,…,M) M components (proteins etc)

dNi/dt= μi Ni → exp(μi t); all μi are equal;

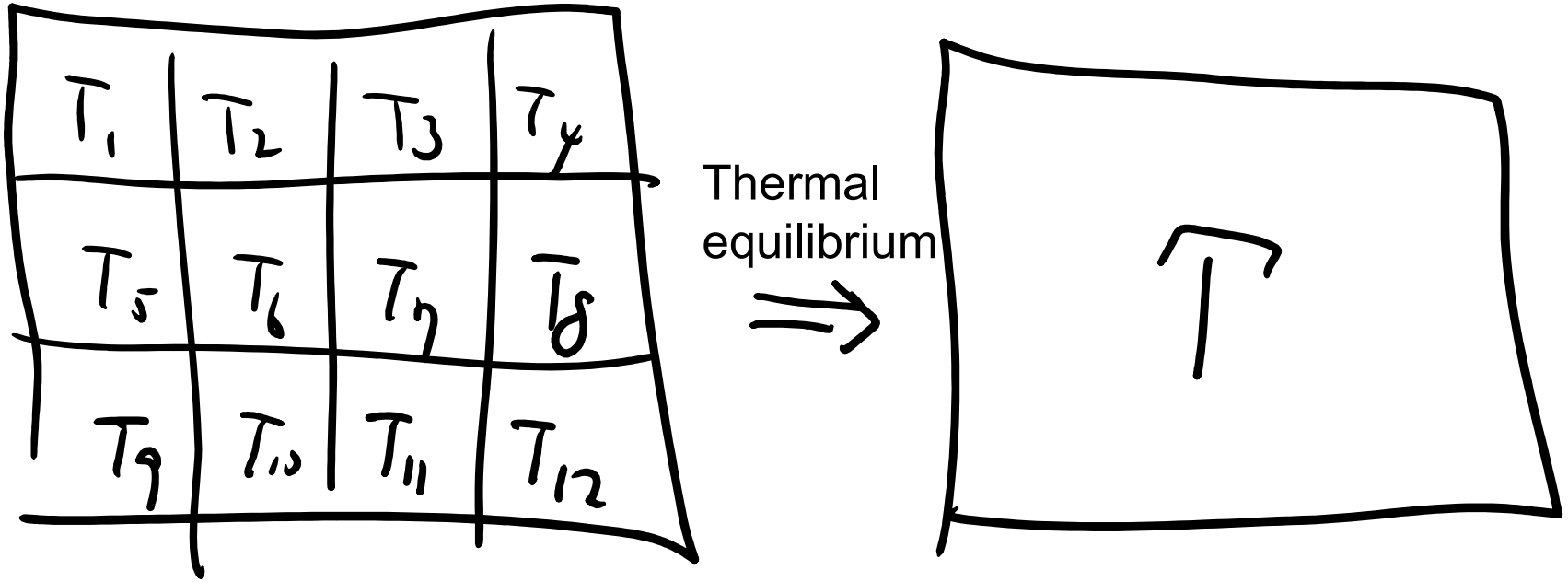
→(M-1) conditions → 1-dimensional line $\sim (10^3 \sim 10^4)$

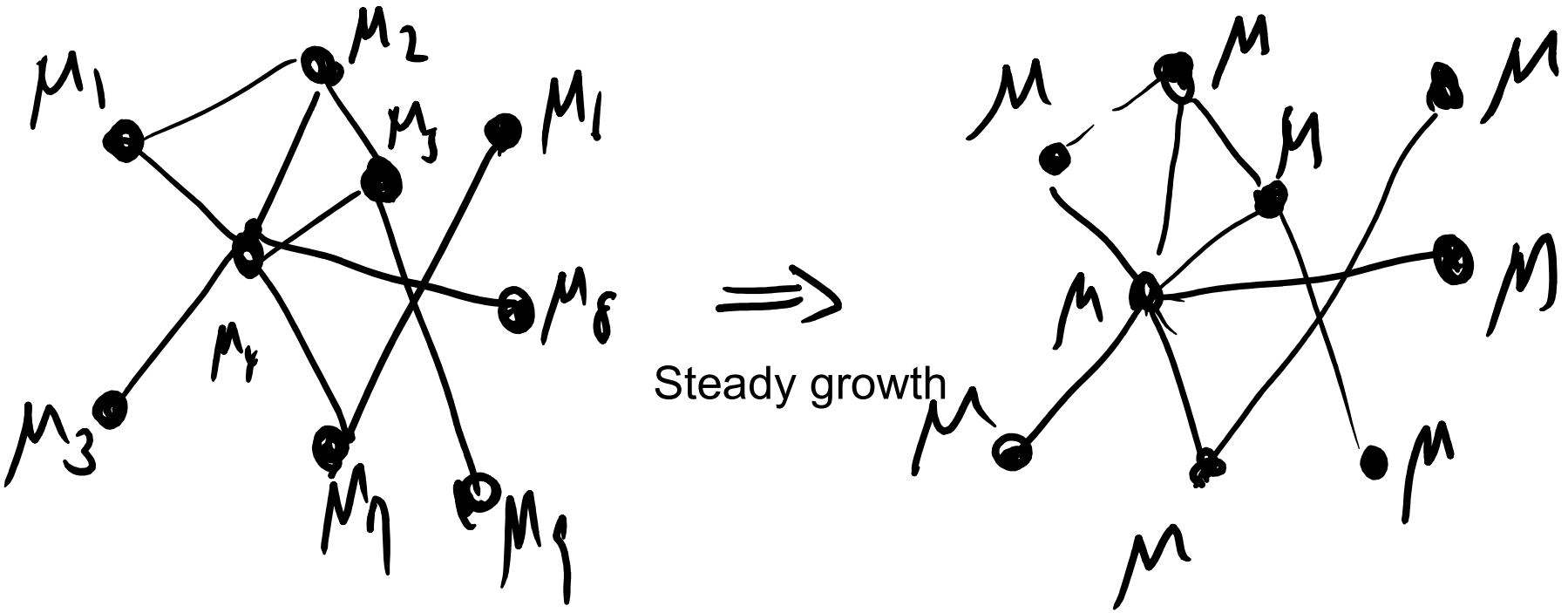


Adaptation on an iso-μi-line
Xi log-concentration, its change δXi, + Linearization at the original state

$$\frac{\delta X_j(E)}{\delta X_j(E')} = \frac{\delta \mu(E)}{\delta \mu(E')} = \text{indep't of j}$$

Restriction to steady growth (here) vs to equilibrium (in thermodynamics): Transient state can involve many degrees





Theory for steady growth: a constraint

Concentration xi=Ni/V: (dV/dt)/V= μ （volume V)

Temporal change of concentration x

$$dx_i/dt = f_i(\{x_j\}) - \mu x_i$$

dilution

fi includes all reactions,
Synthesis, degradation,..

Now, the stationary state is given by a fixed point condition

$$x_i^* = f_i(\{x_j^*\})/\mu$$

for all i.

As a convenience, denote $X = logx$, and $f_i = x_i F_i$. Then,

$$dX_i/dt = F_i(\{X_j\}) - \mu$$

Response under different stress strength E

$$F_i(\{X_i^*(E)\}, E) = \mu(E).$$

Linearization around original statew.r.t X(=log x)

KK, Furusawa Yomo.
Phys Rev X(2015)

$$\sum_j J_{ij}\delta X_j(E) + \gamma_i \delta E = \delta\mu(E)$$

Jacobi matrix J_{ij}.

with $\gamma_i \equiv \frac{\partial F_i}{\partial E}$. ← Susceptibility to stress

In the linear regime $\delta\mu = \alpha\delta E.$

$$\delta X_j(E) = \delta\mu(E) \times \sum_i L_{ji}(1 - \gamma_i/\alpha)$$ $L = J^{-1}.$

$$\frac{\delta X_j(E)}{\delta X_j(E')} = \frac{\delta\mu(E)}{\delta\mu(E')}$$ = indep't of j

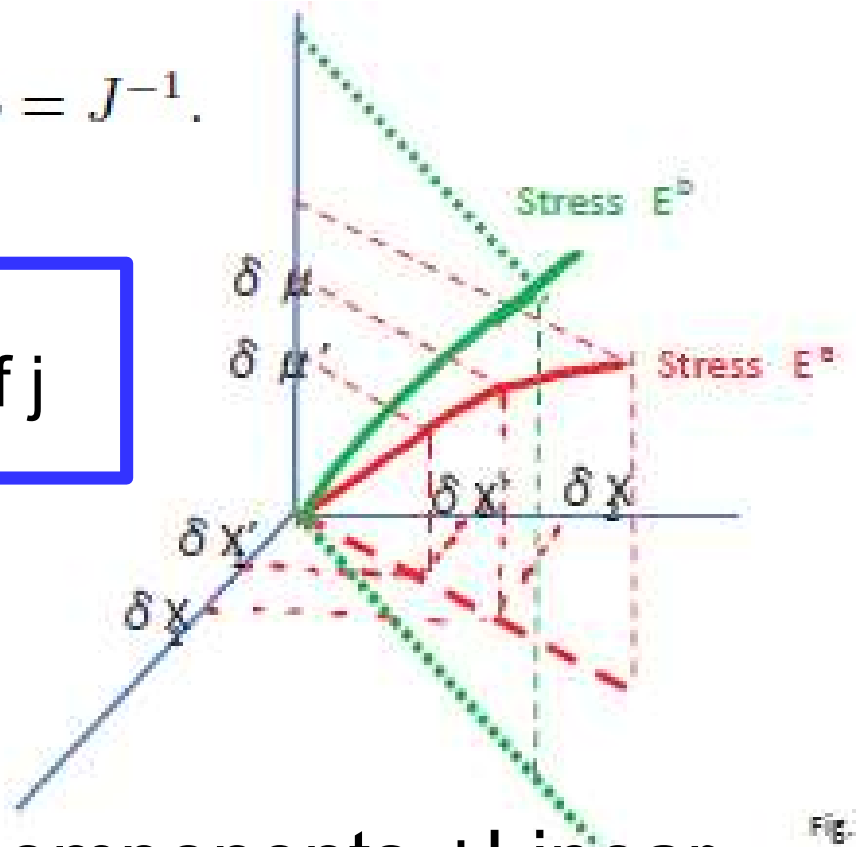


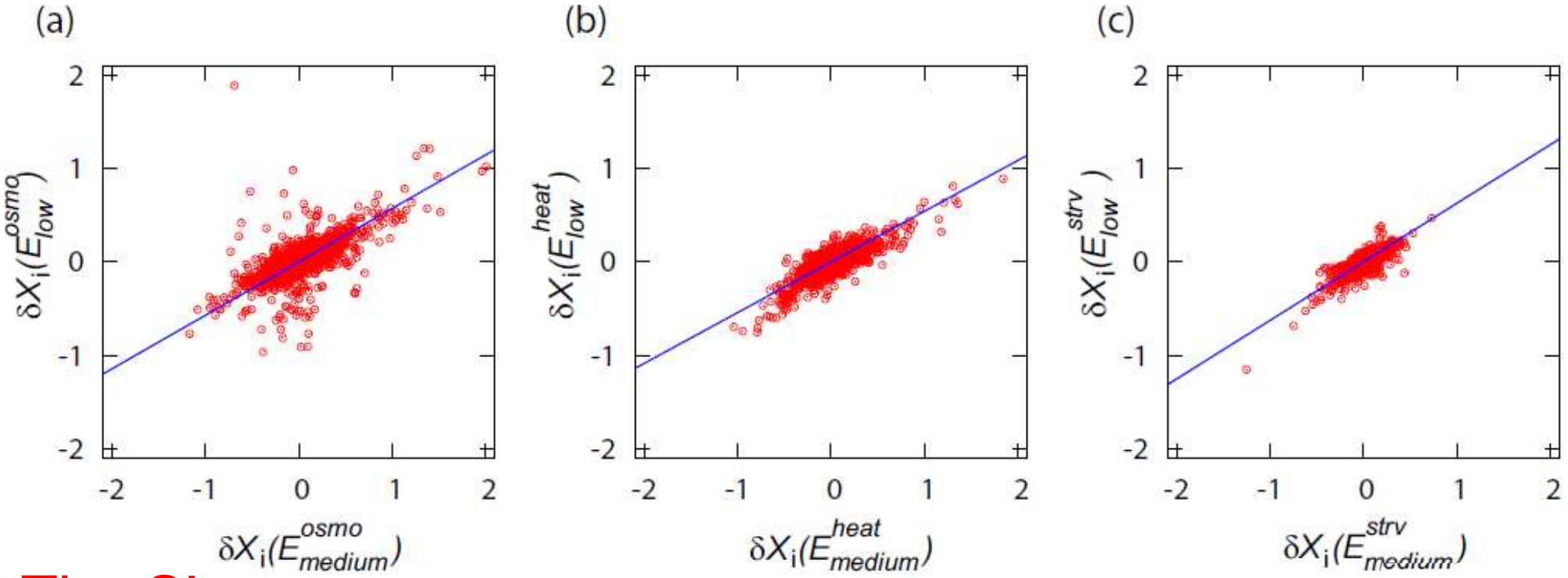
Fig. 2b

Common proportionality for log-expression change δXj for all components j

←Steady-growth sustaining all components +Linear

Put E Coli under different strength of stress conditions; Measure gene expressions

$$log(x_i(E)/x_i^O) \text{ and } log(x_i(E')/x_i^O)$$

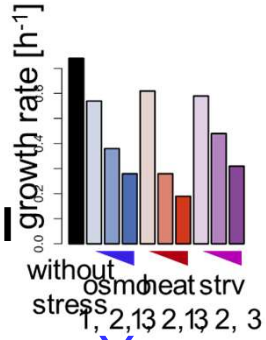


The Slope agrees with the growth rate change
δμ'/δμ

A: low vs medium osmo
B low vs medium heat
C low vs medium starvation

δX^E、δX^E'
over few thousand genes

Data from Matsumoto etal BMC Evol Biol I2013



KK,Furusawa,Yomo, Phys Rev X (2015)

Compare between Original state O and the states at stresses E or E'

$$log(x_i(E)/x_i^O) \text{ and } log(x_i(E')/x_i^O)$$

Assuming that cellular states are stationary (growth-rates of all components are balanced)
"quasi-stationary-processe

On the average the growth rates are balanced

(if oscillatory, take temporal average)

＋ No Bifurcation to different branches of solution in F(X)

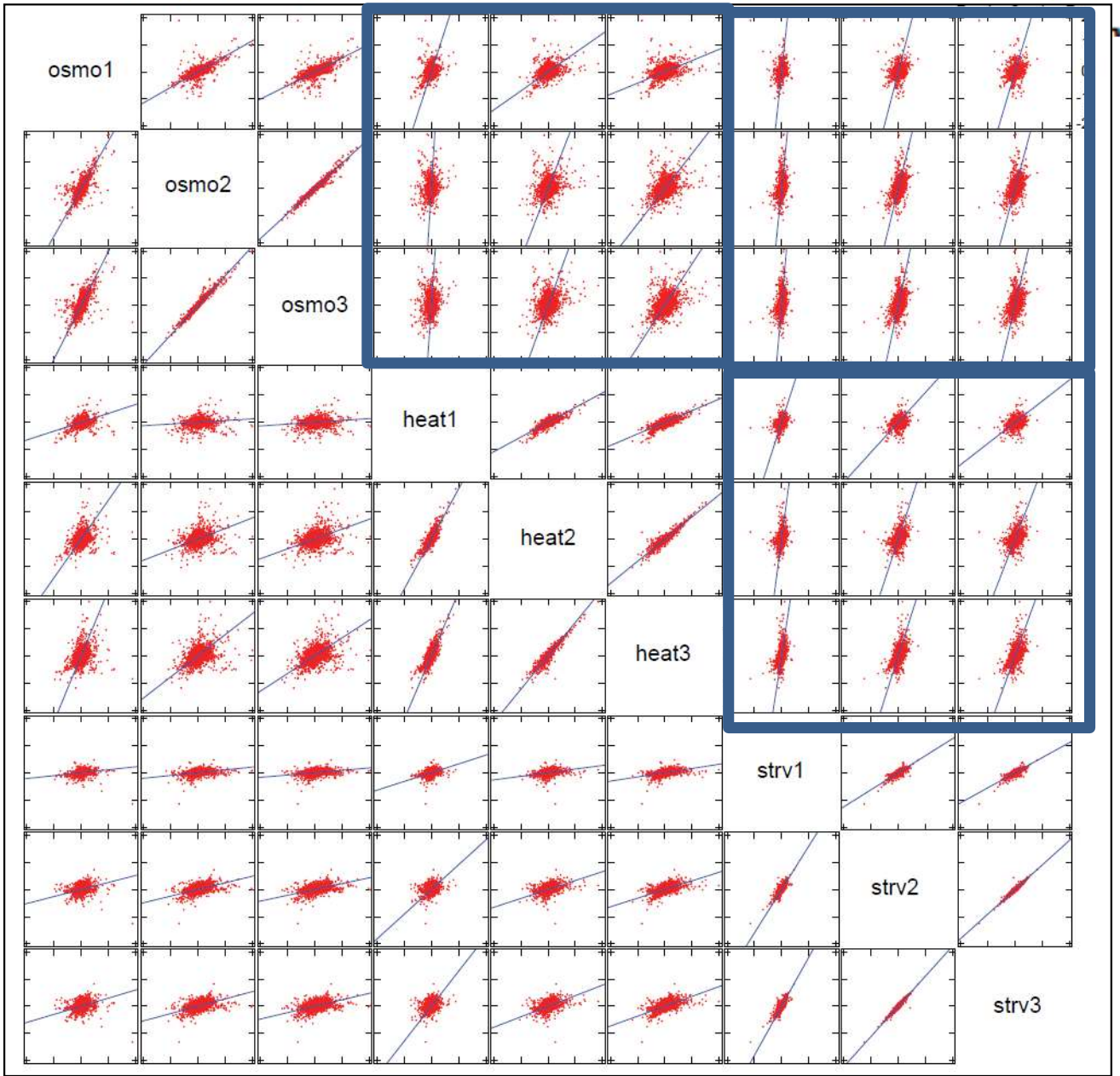
＋ Linearization in X （log x）

Transcriptome expreiment

Comparison across different stress conditions;

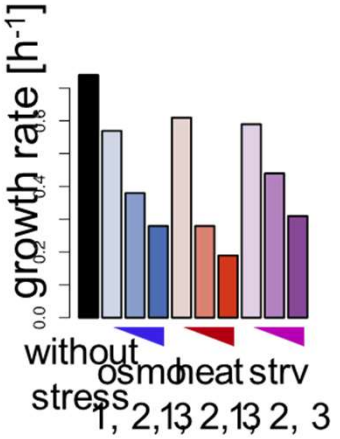
$$log(x_i(E)/x_i^O) \text{ and } log(x_i(E')/x_i^O)$$

expressions



O: no stress
E,E' : osmotic pressure, heat, starvation
Low, medium high

Matsumoto,Yomo
NatCellBiol2013



Compare Different types of stresses: Jacobi matrix J_{ij}.

$$\sum_j J_{ij}\delta X_j(E) + \gamma_i \delta E = \delta\mu(E)$$

with $\gamma_i \equiv \frac{\partial F_i}{\partial E}$. →γi depends on stress type (*a,b,..*)

$\delta\mu = \alpha\delta E.$ →α depends on stress type (*a,b,..*)

$$\delta X_j(E) = \delta\mu(E) \times \sum_i L_{ji}(1 - \gamma_i/\alpha)$$

$L = J^{-1}$.

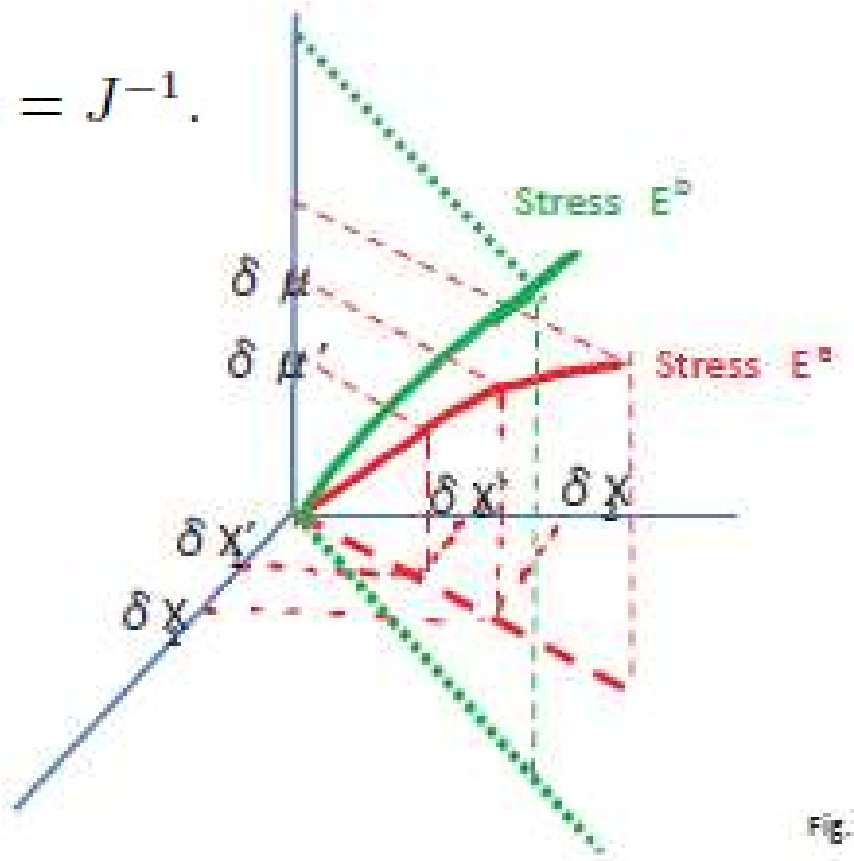


Fig. 2b

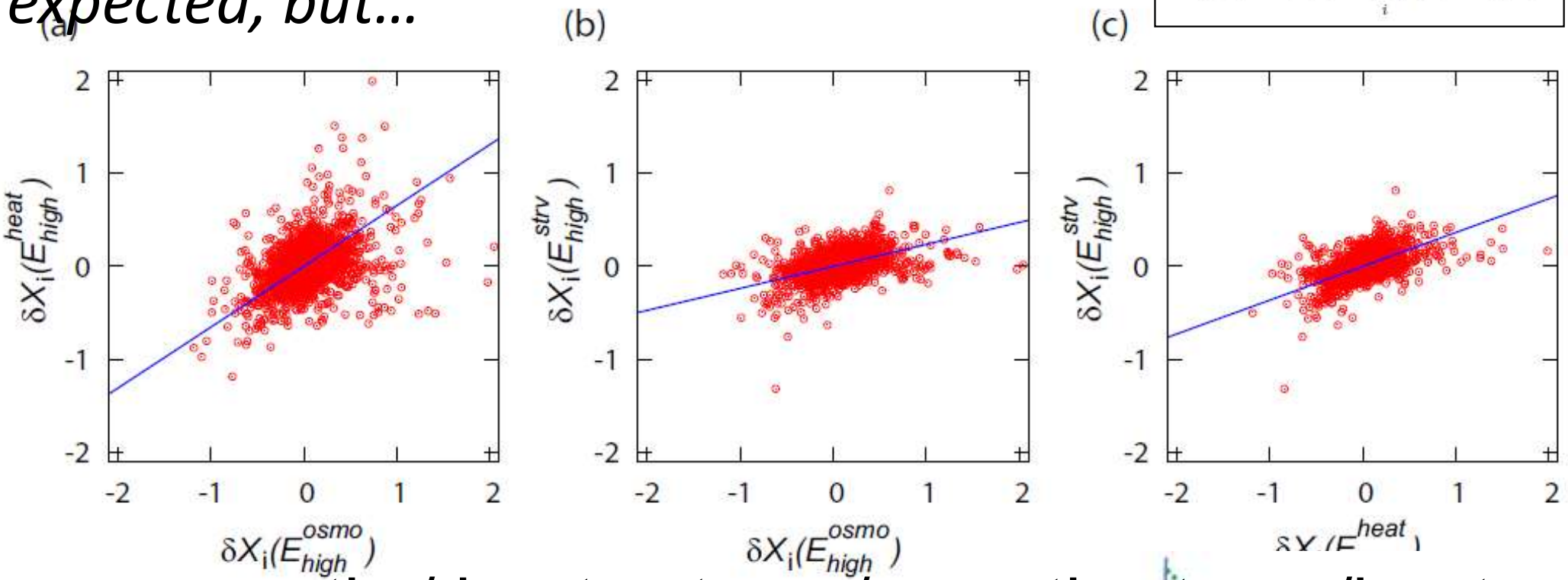
$$\frac{\delta X_j(E)}{\delta X_j(E')} = \frac{\delta\mu(E)}{\delta\mu(E')}$$

Across Different types of stresses:

$$\gamma_i \equiv \frac{\partial F_i}{\partial E}$$

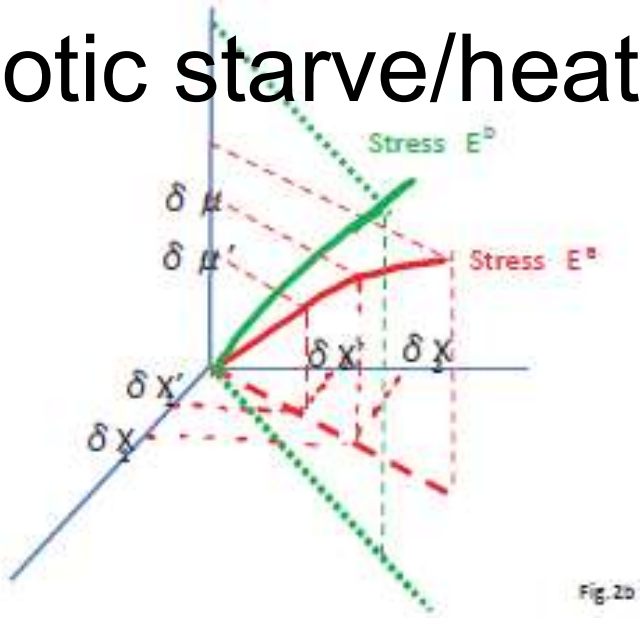
γi(*a*) depends on type *a so correlation might not be expected, but…*

$$\delta X_j(E) = \delta\mu(E) \times \sum_i L_{ji}(1 - \gamma_i/\alpha)$$

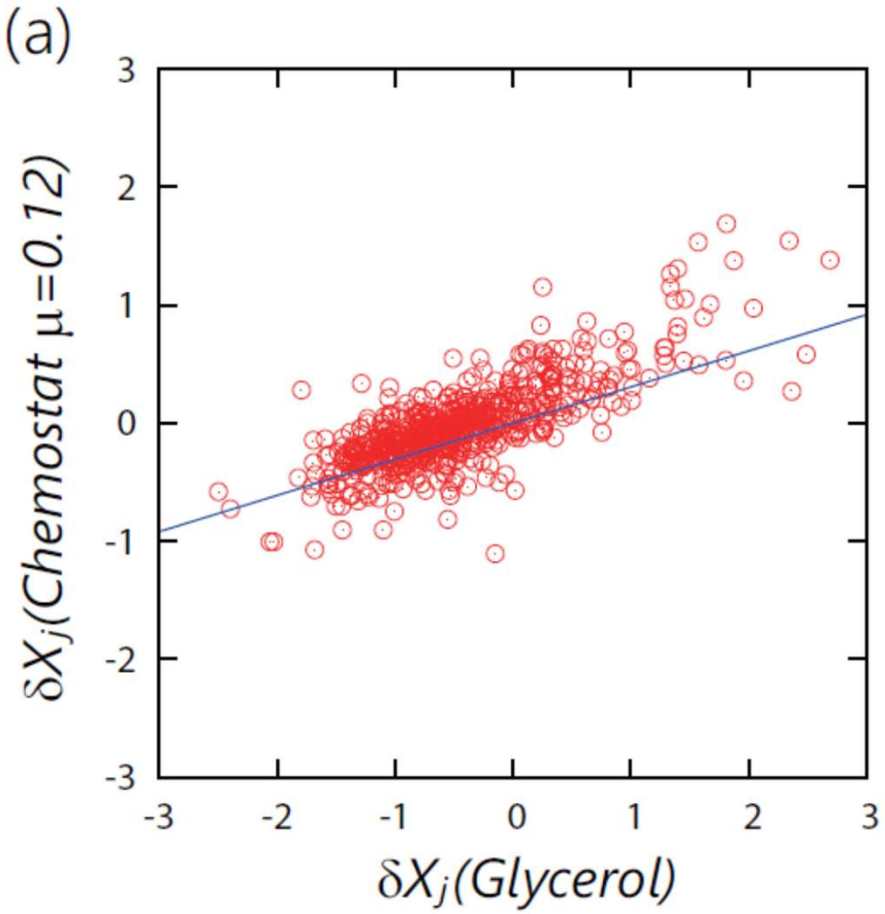


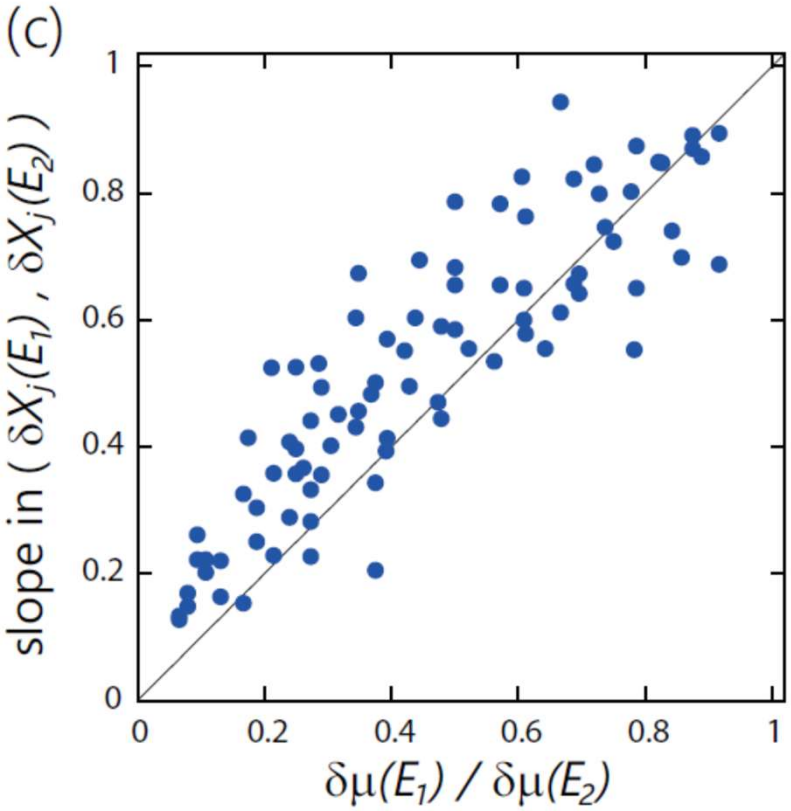
osmotic / heat starve/osmotic starve/heat

Still highly correlated



Confirmed also protein expression changes across different environmental conditions (based on the data by Heinemann) 20 different conditions on E Coli





Furusawa, KK bioRxiv 2017

Non-trivial point: Emergent macroscopic Linearity

- (1) Large Linear Regime?
- (2) Validity across different environmental condition?

Q : achieved in an evolved system(to macro regime)?

before addressing it..

Is this universal relationship extended to evolution-environmental relationship? ← LATER

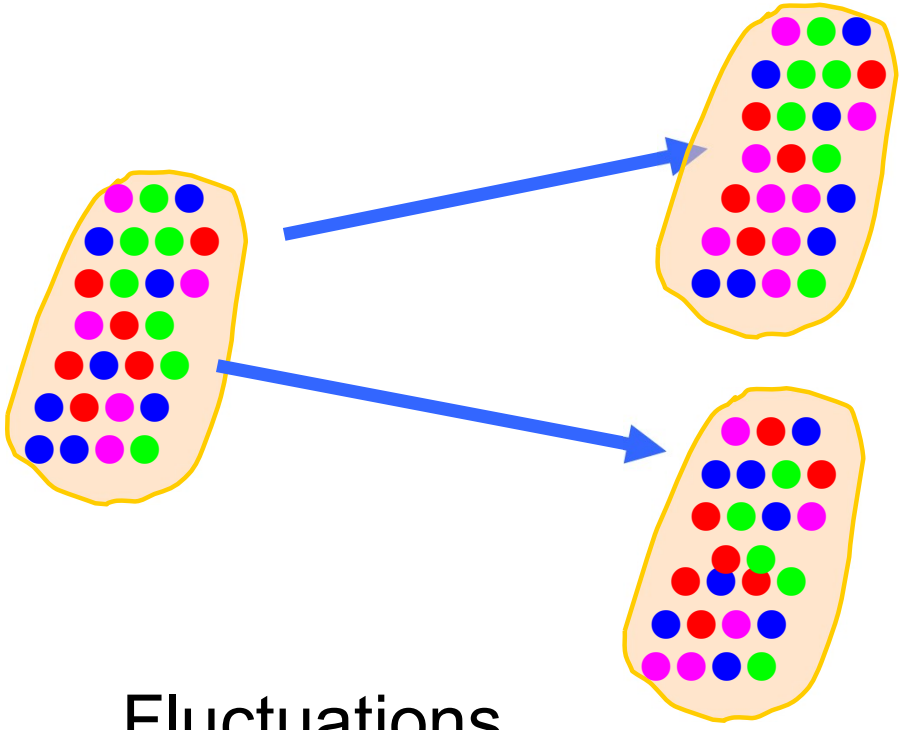
How is recursive production of a cell sustained?

each cell complex reaction network

with diversity of chemicals;

The number of molecules of each species

not so large



?

Fluctuations

Naiive Physicist View

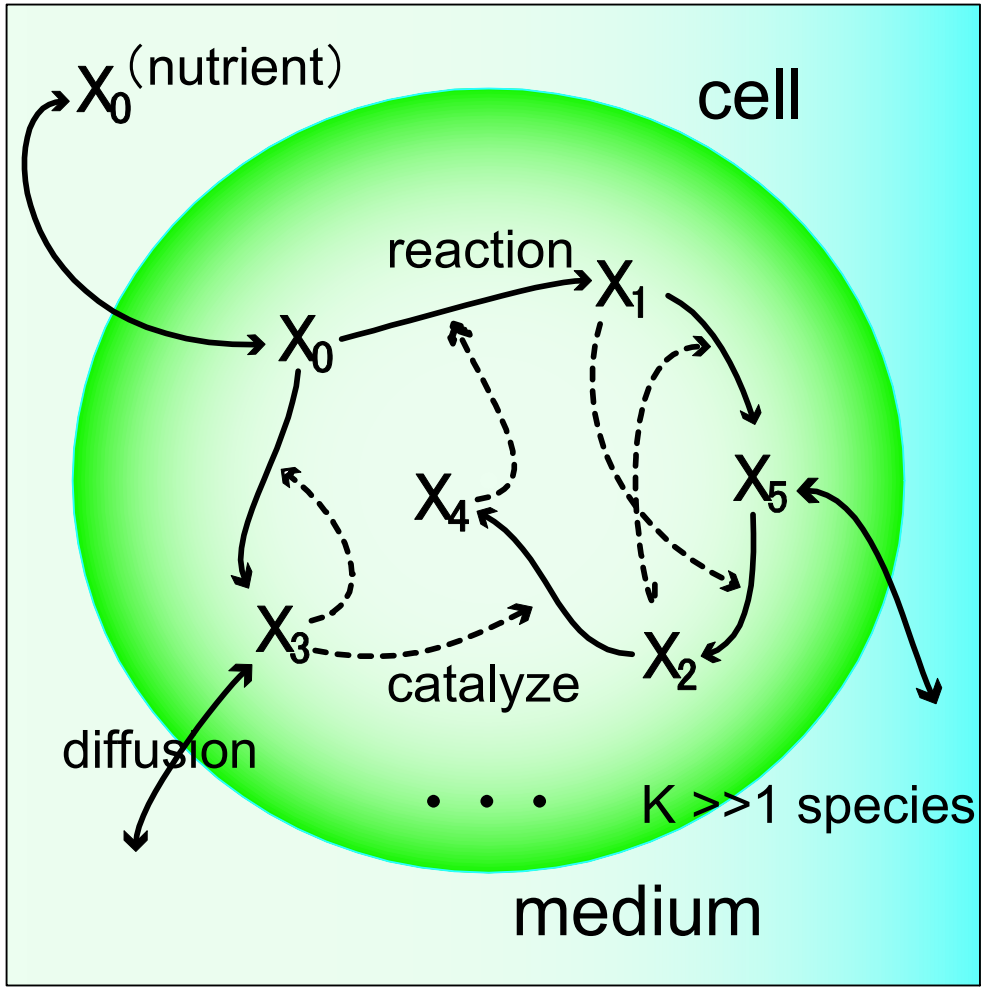
Not Assuming Molecule Replication (Replicators)

Toy Cell Model with Catalytic Reaction Network
'Crude but whole cell model'

C.Furusawa & KK、PRL2003

- **k species of chemicals** 、$X_o \cdots X_{k-1}$ number ---n_0 、$n_1 \ldots n_{k-1}$
- random catalytic reaction network
  with the path rate p
  for the reaction $X_i + X_j -> X_k + X_j$
- some chemicals are penetrable through the membrane with the diffusion coefficient D
- resource chemicals are thus transformed into impenetrable chemicals, leading to the growth in $N = \Sigma n_i$, when it exceeds N_{max} the cell divides into two

model (Cf. KK&Yomo 94,97)



dX1/dt ∝ X0X4; rate equation;
Stochastic model here

☆Simulation procedure

1 : Pick up randomly 2 molecules at each time step, if the pair reactions, change the substratre molecule into productt (with the probability of the reaction rate) otherwise leave as it is

2 With a certain rate per time step (（≈1／D）, exchange a molecule of inside in the cell by that in an enviromenment. If the molecule is impermeable, it stays

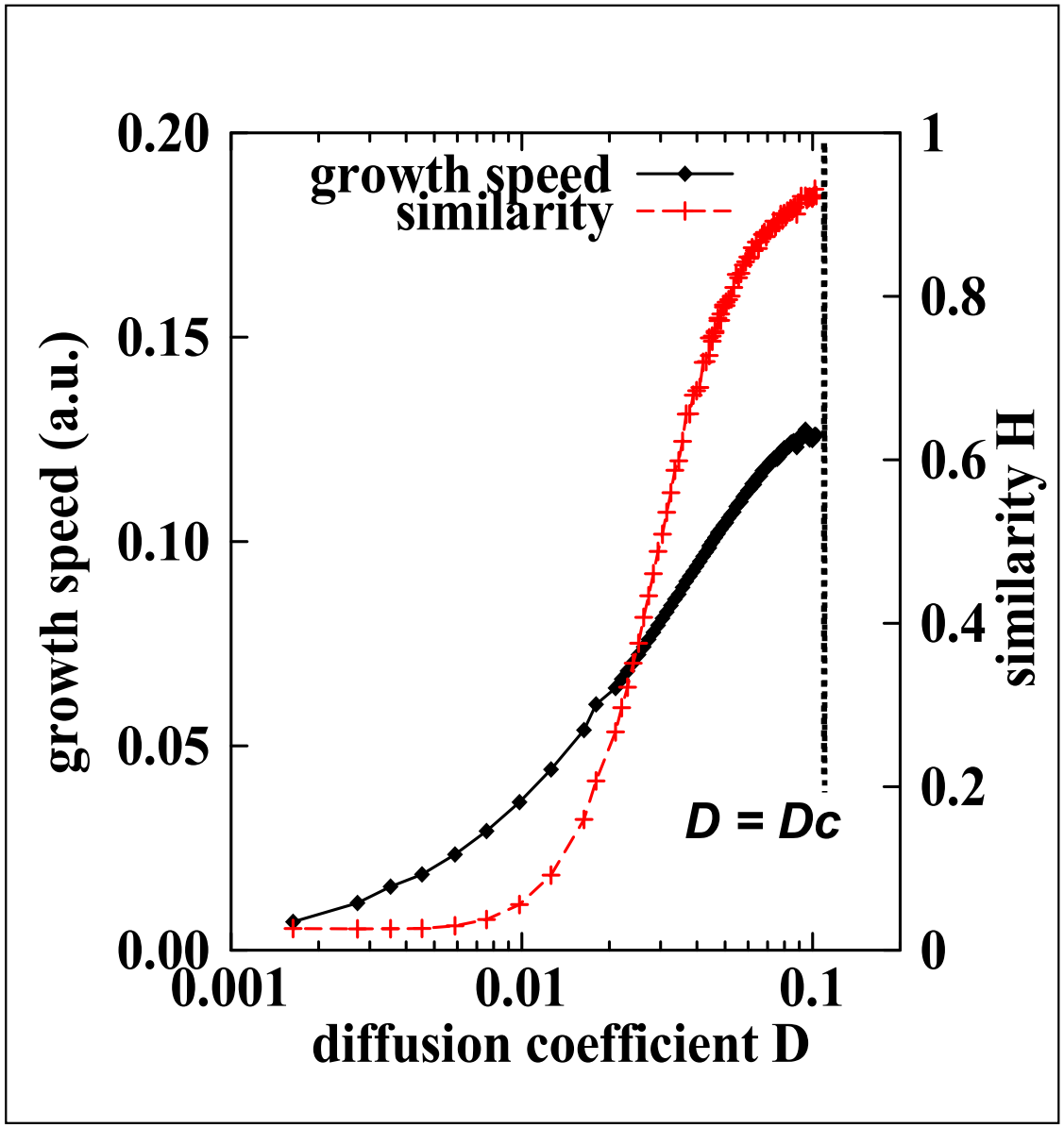
3 : If the total number of molecules N goes beyodn N_{max} cells are divided into two, eahc of which consists of molecules chosen randomly

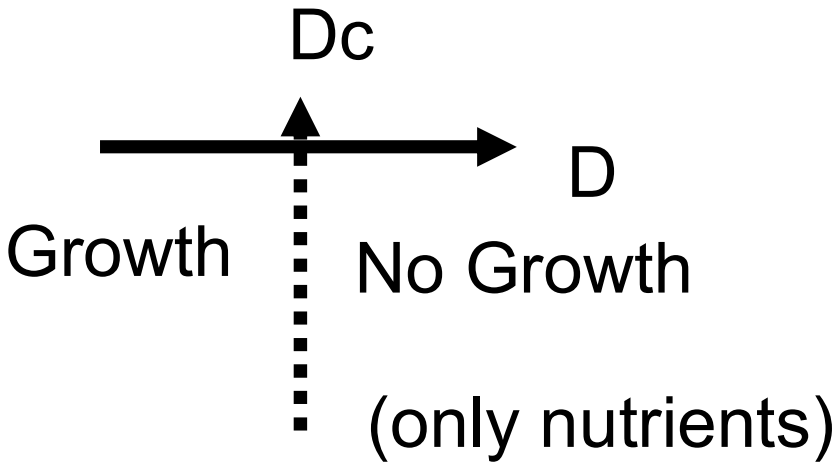
In continuum description, the following rate eqn., but we mostly use stochastic simulation

$$dn_i/dt = \sum_{j,\ell} \mathrm{Con}(j, i, \ell)\epsilon n_j n_\ell / N^2 - \sum_{j',\ell'} \mathrm{Con}(i, j', \ell')\epsilon n_i n_{\ell'} / N^2 + D\sigma_i(\overline{n}_i/V - n_i/N),$$

where $\mathrm{Con}(i, j, \ell)$ is 1 if there is a reaction $i + \ell \rightarrow j + \ell$, and 0 otherwise, whereas σ_i takes 1 if the chemical i is penetrable, and 0 otherwise. The third term describes the transport of chemicals through the membrane, where $\overline{n}_i$ is

☆Growth speed and fidelity in replication are maximum at Dc

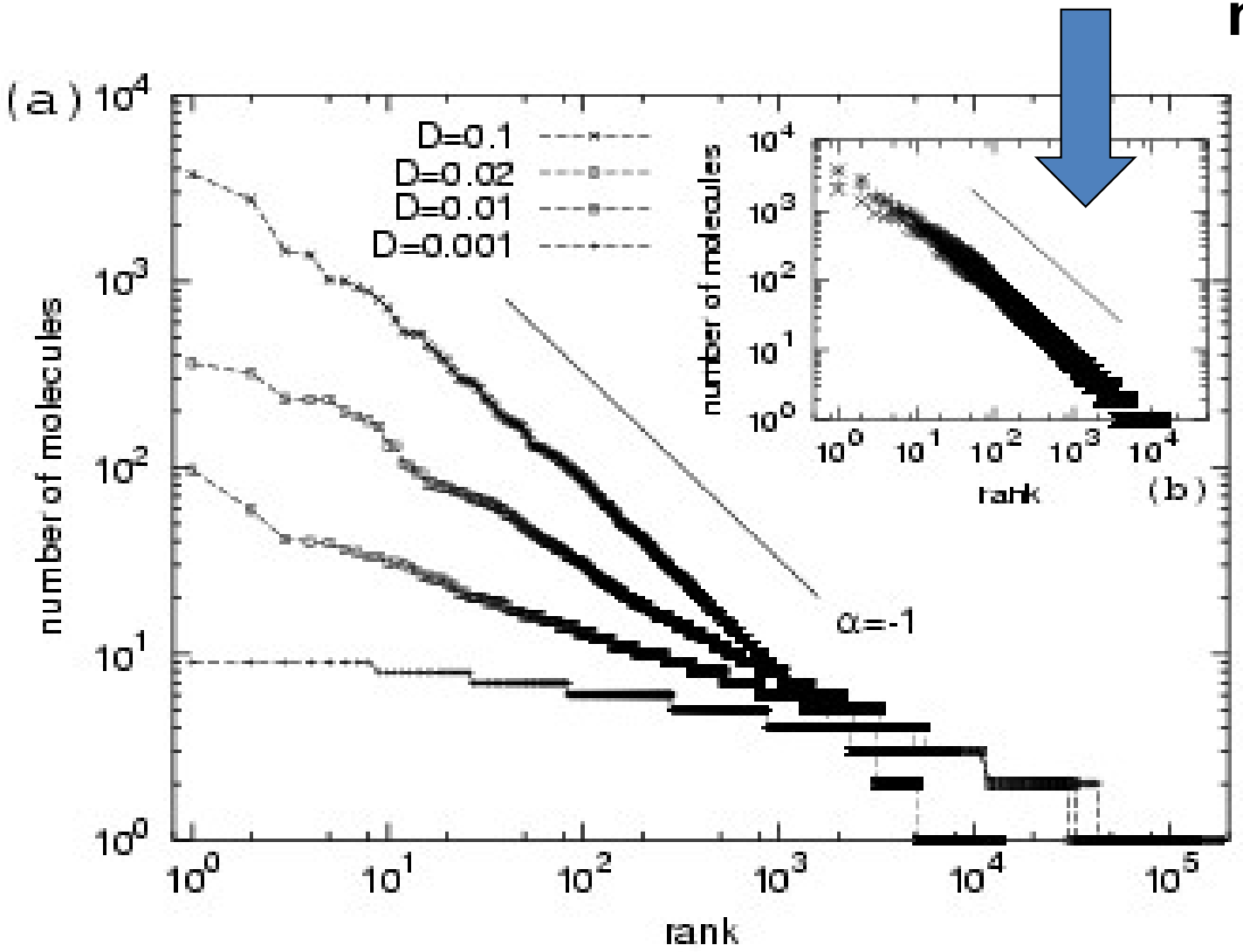




※similarity is defined from inner products of composition vectors between mother and daughter cells

Furusawa &KK,2003,PRL

Zipf's Law is oberved at D = Dc



n_i (number of molecules)

rank

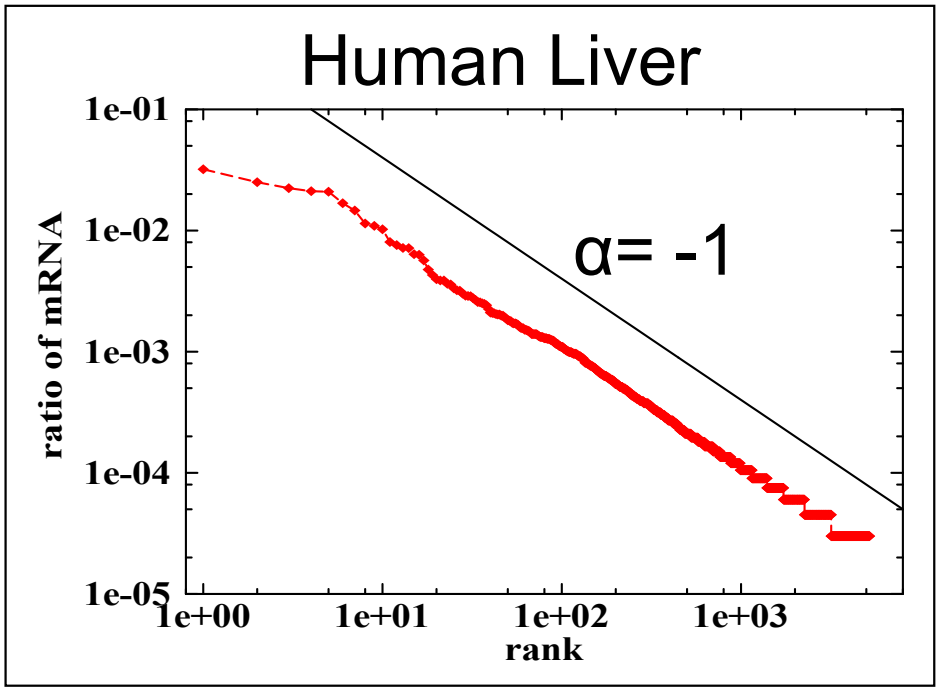
	number	rank
X1	300	5
X2	8000	1
X3	5000	2
X4	700	4
X5	2000	3

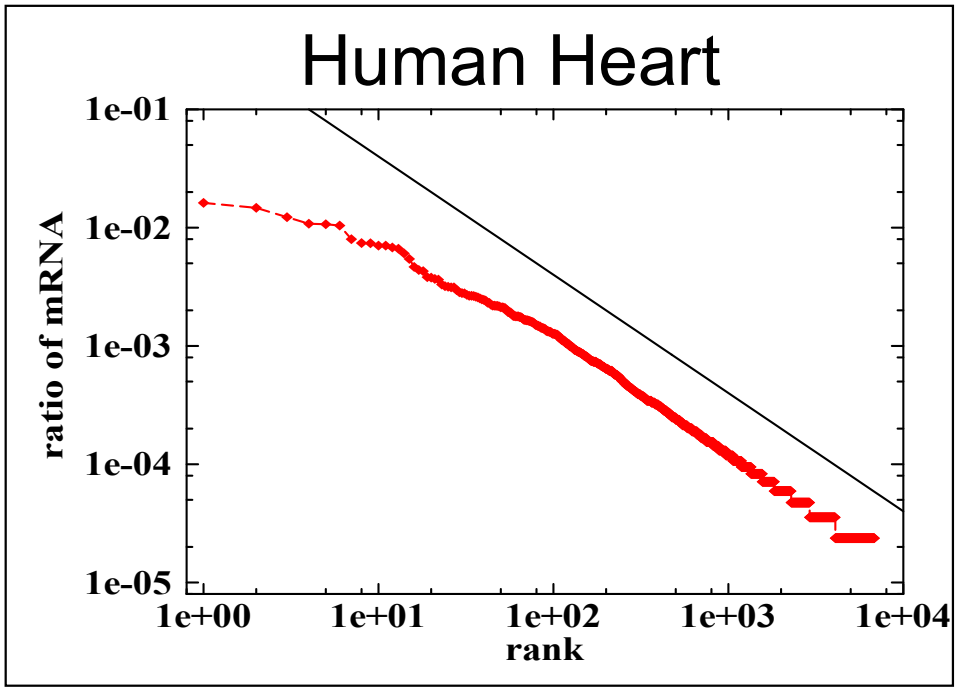
…….. (for example)

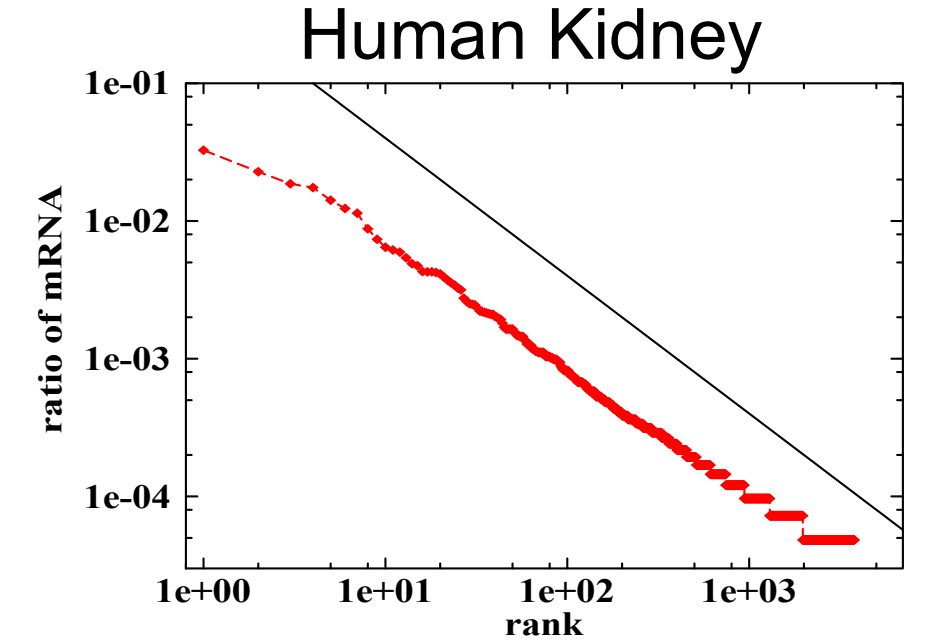
Average number of each chemical ∝ 1/(its rank)

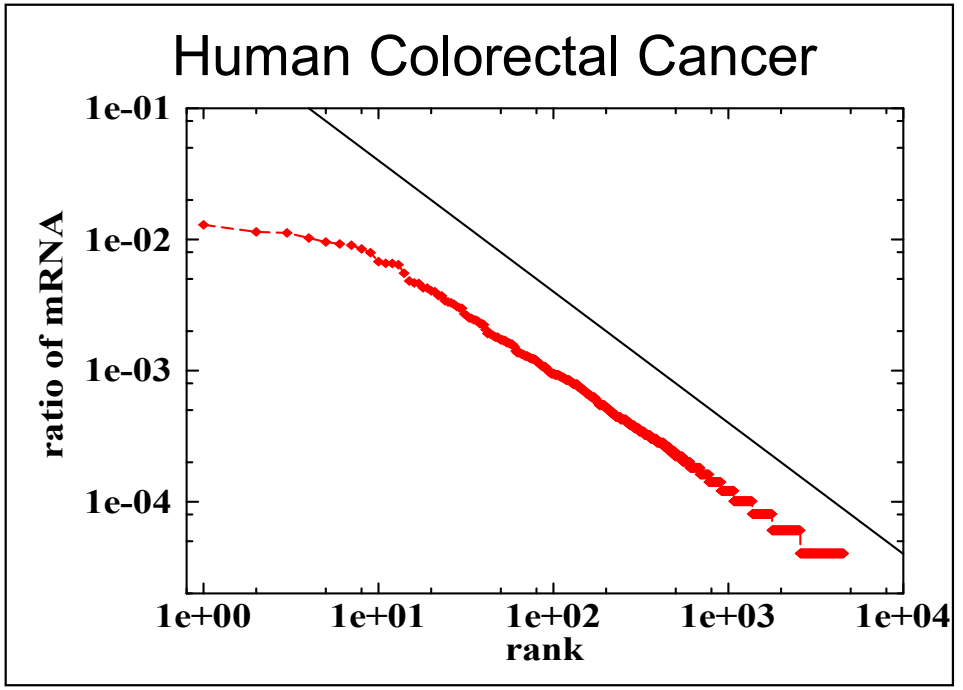
(distribution of $x : \rho(\mathbf{x}) \propto \mathbf{x}^{-2}$)

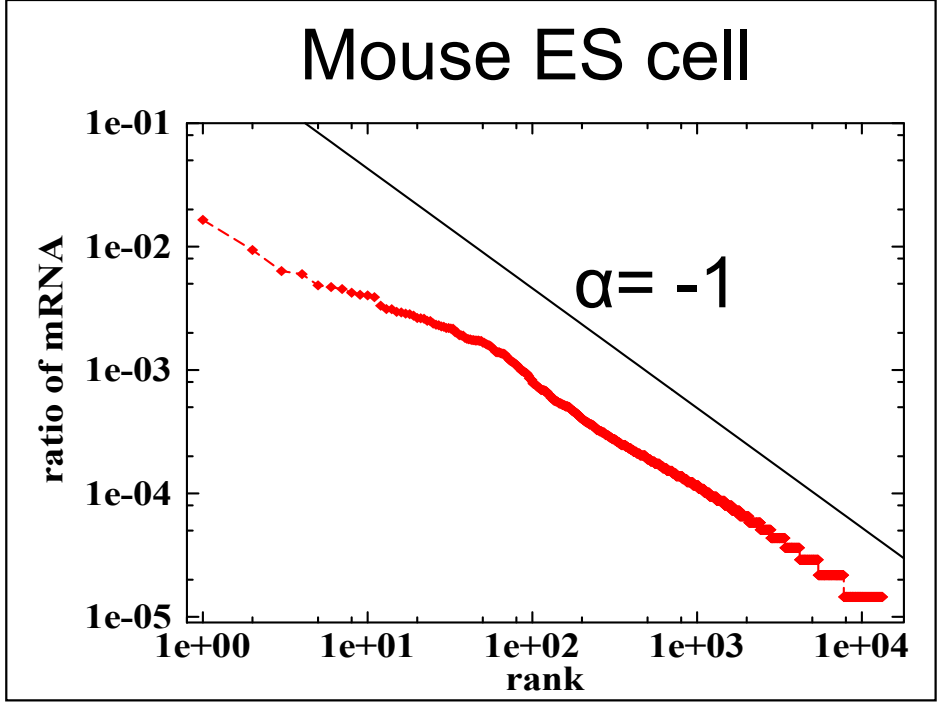
Confirmed by gene expression data

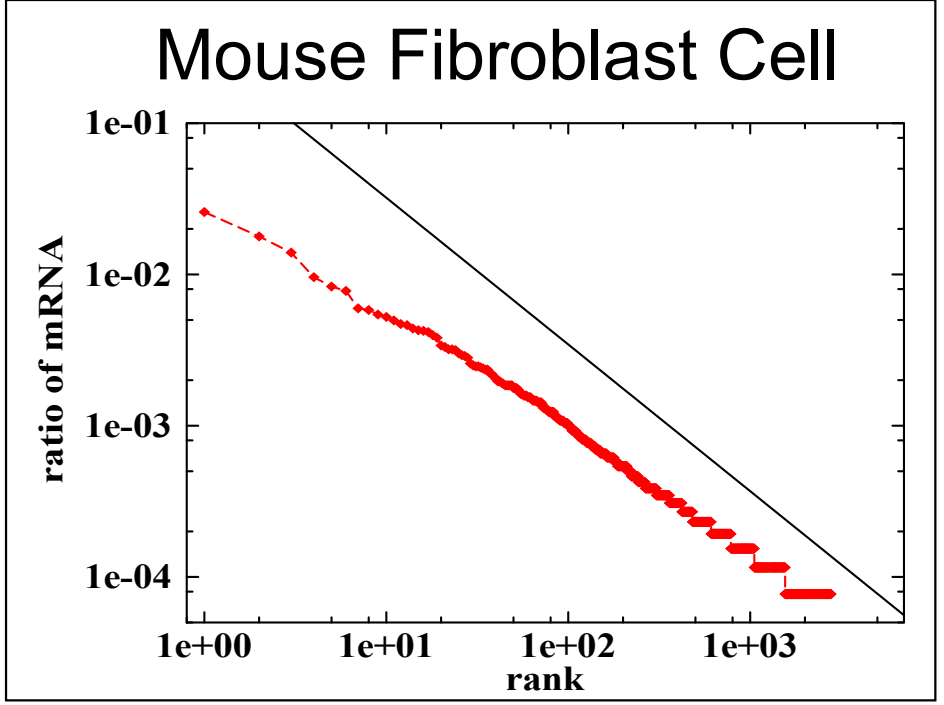


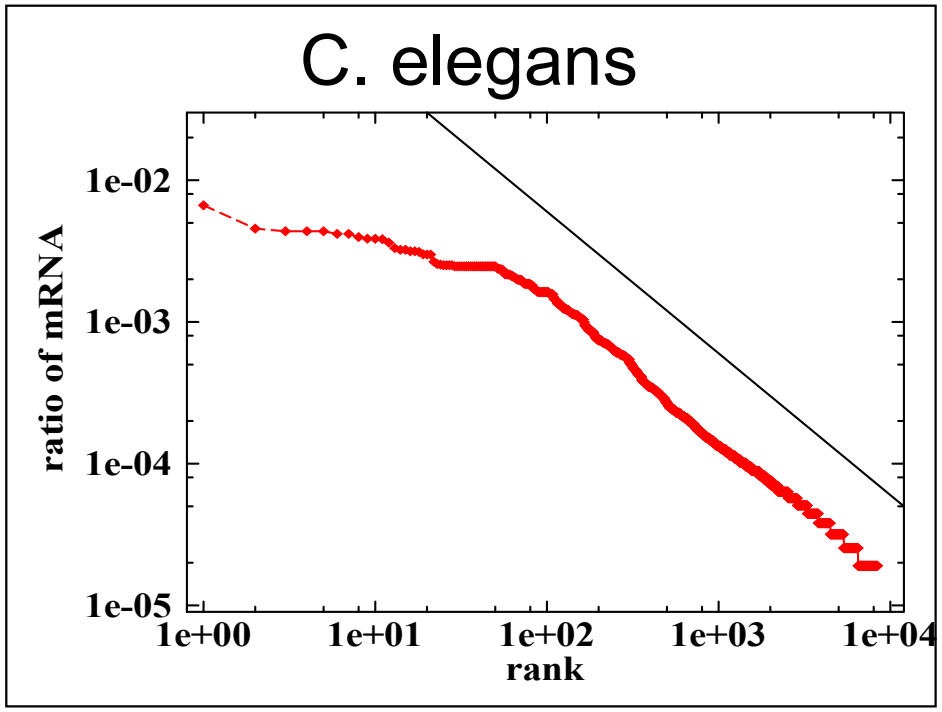


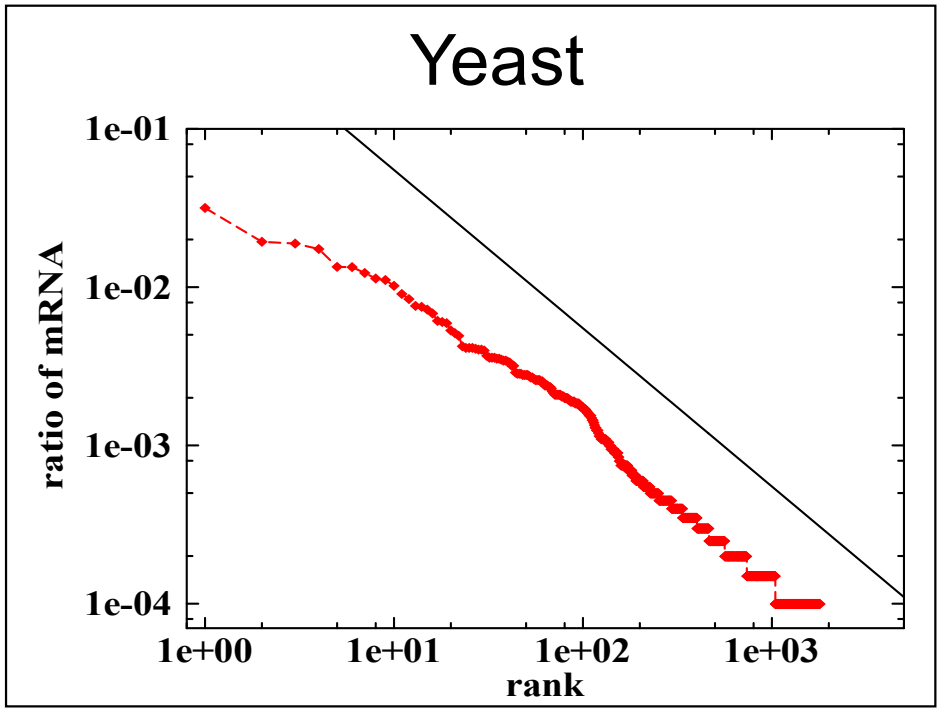






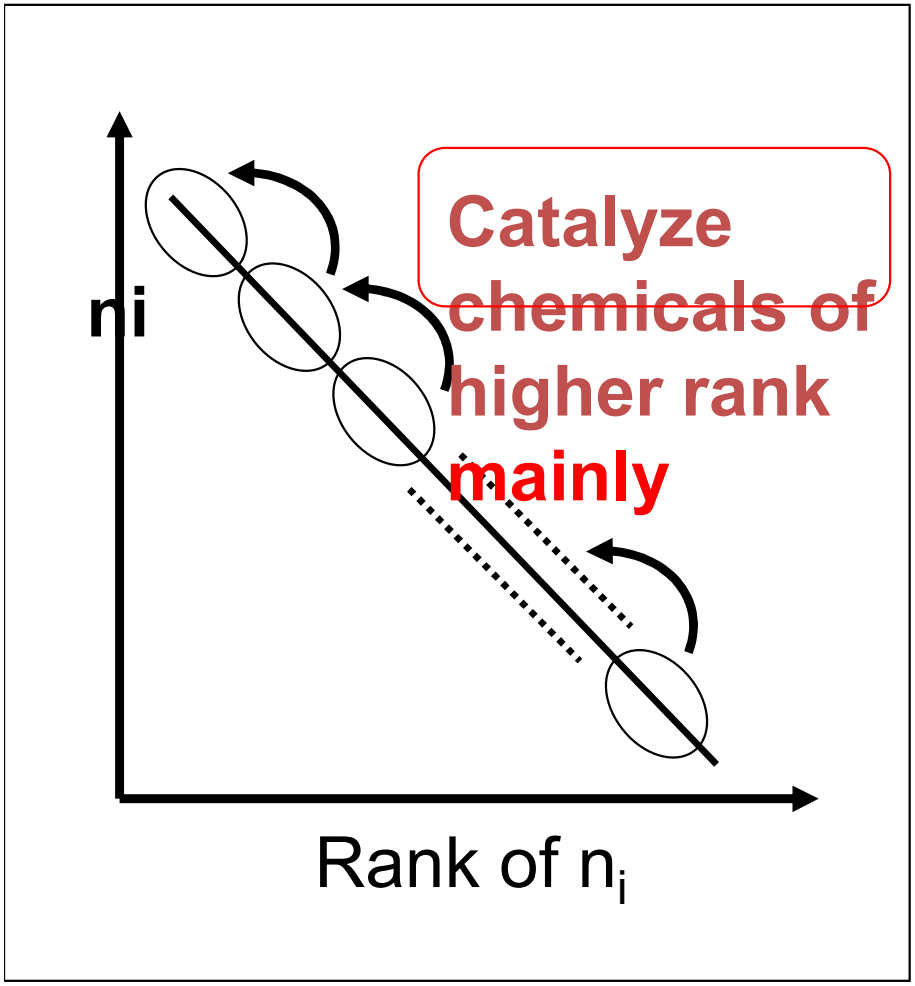




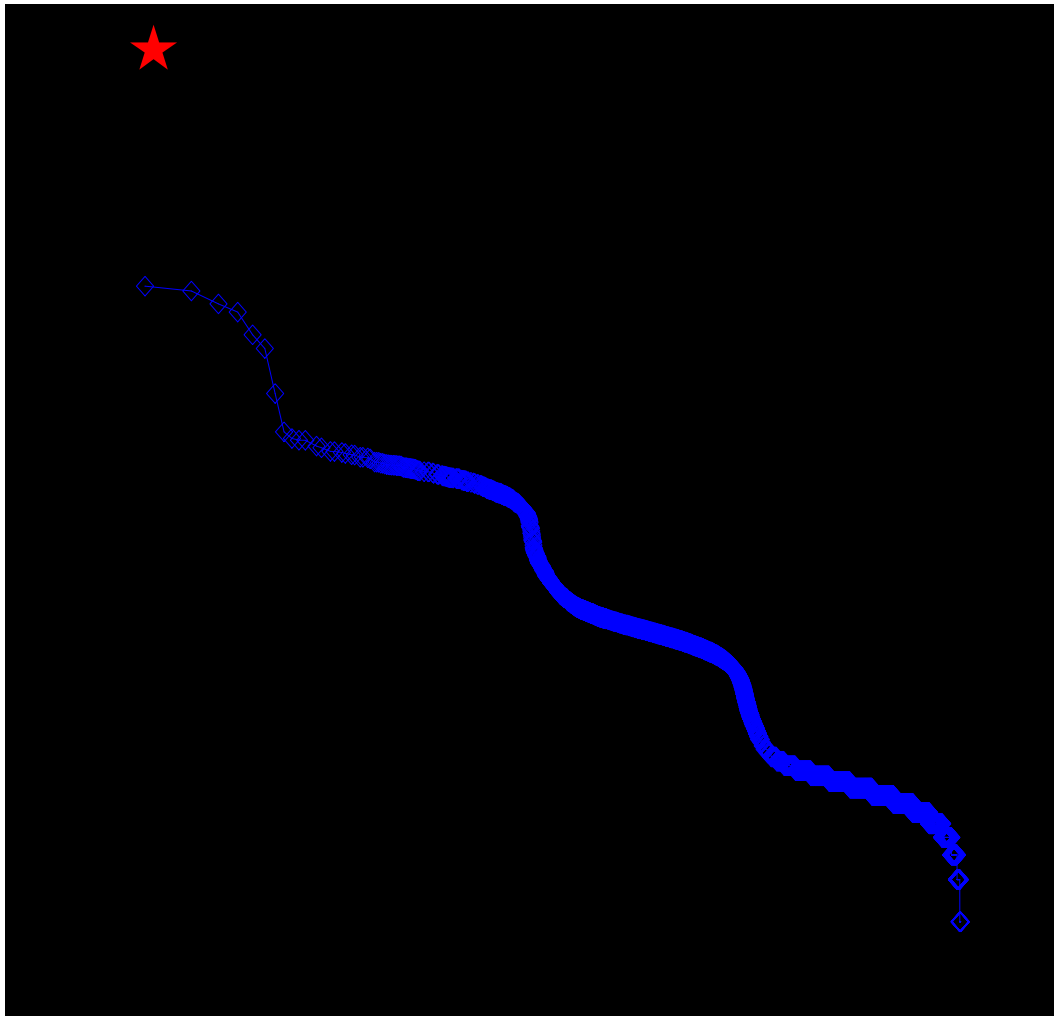


Later confirmed by several other groups

Formation of cascade catalytic reaction



With conservation law,
The exponent -1 is explained



1 : minority molecules

2 : catalyzed by 1, synthesized by resource

3 : catalyzed by 2

Mean-field theory in phase transition (self-consistent) calc.)

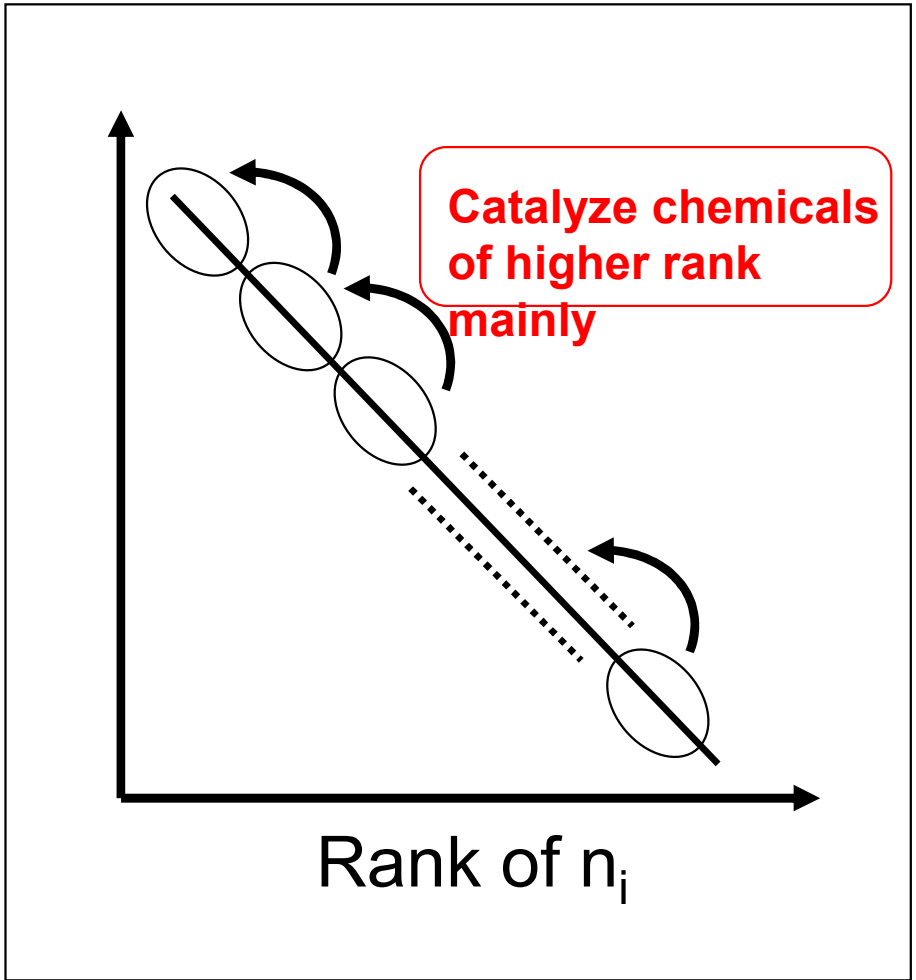
- Simple laws hold in real biological organisms

- The abundance-ranking inverse law is often observed

  frequency of words （the and of…） Zipf's law
  ranking of income

Successive ordering in mltual catalytic reactions

Scale invariance,
Phase transition



Fluctuation of each chemical Abundance;

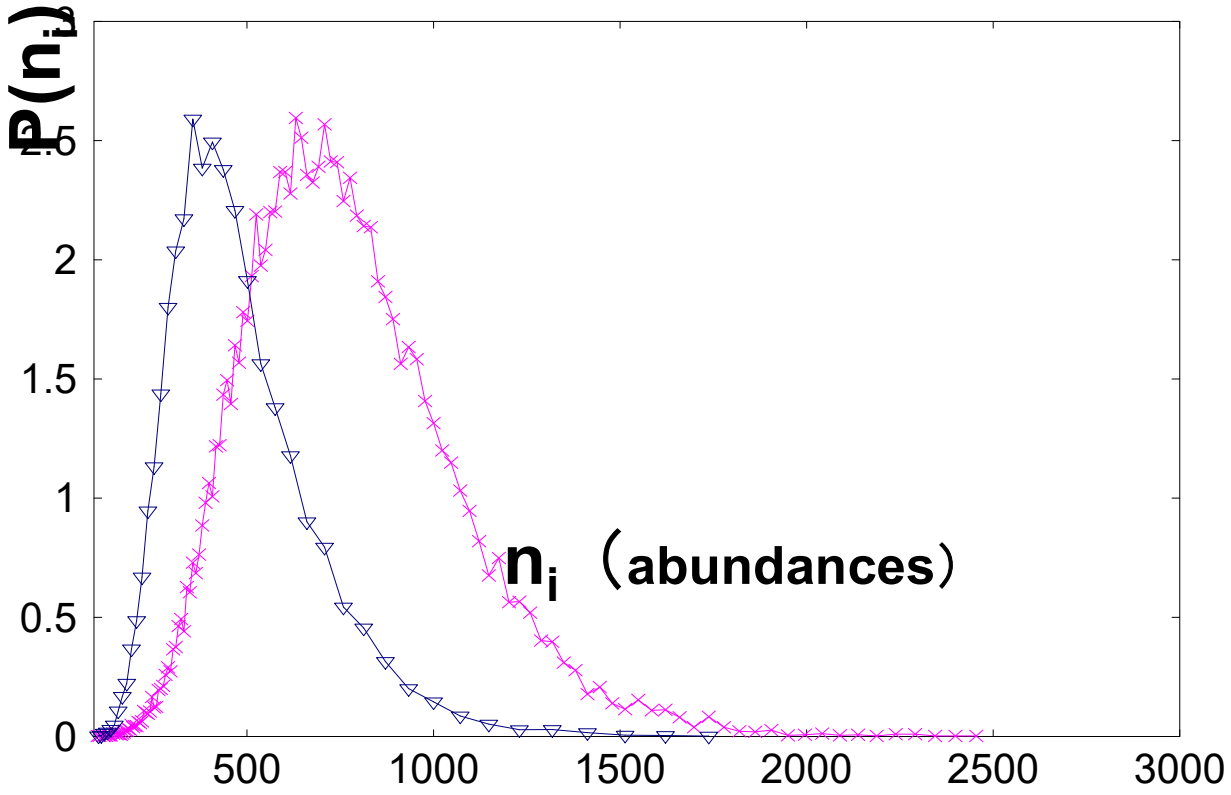
→ long-tail to abundant size

e.g.

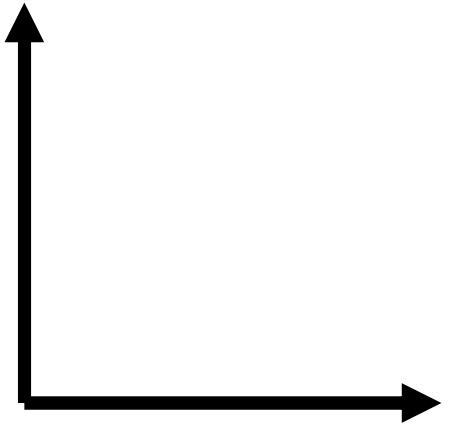
cell1 X1	10000
cell2	8000
cell3	15000
cell4	20000

…..

histogrram



Frequency of n_i

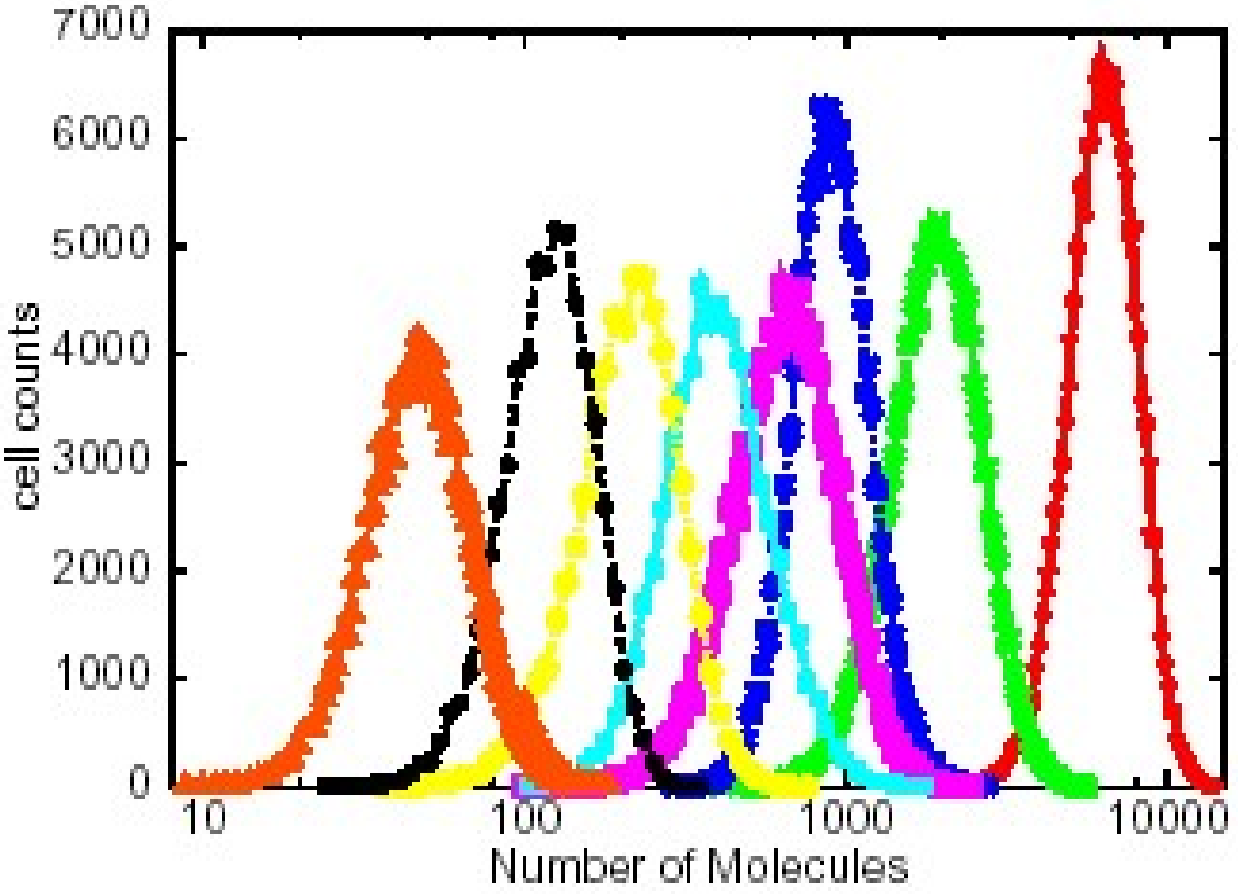


n_i (abundances)

So far average quantity of all components;

Next question: fluctuation by cells:
distribution of each Ni by cells

Log normal distribution !



Furusawa,..
KK,
Biophysics2005

e.g.

cell1 X1	10000
cell2	8000
cell3	15000
cell4	20000

.....

histograrm

Each color gives different chemical species

LOG SCALE

☆Heuristic explanation of log-normal distribution

Consider the case that a component X is catalyzed by other component A, and replicate; the number --N_X、N_A

$$d N_X /dt = N_X N_A$$

then

$$d \log(N_X)/dt = N_A$$

If、N_A fluctuates around its mean $<N_A>$, **with fluct.** η (t)

$$d \log(N_X)/dt = <N_A> + \eta (t)$$

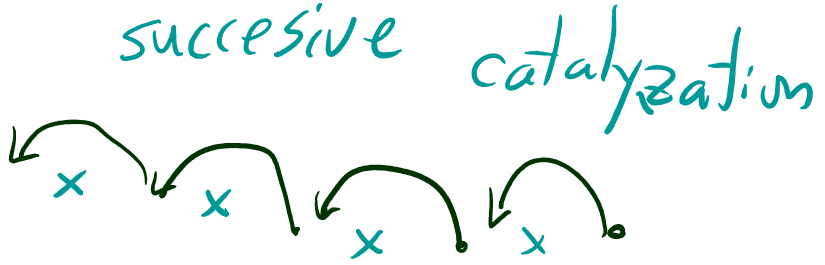
log(N_X) shows Brownian motion → N_X log-normal distribution

too, simplified, since no direct self-replication exists here

But with cascade catalytic reactions, fluctuations are successively multiplied, (cf addition in central limit theorem.);Hence after logarithm, central limit th. applied

☆Heuristic explanation of log-normal distribution

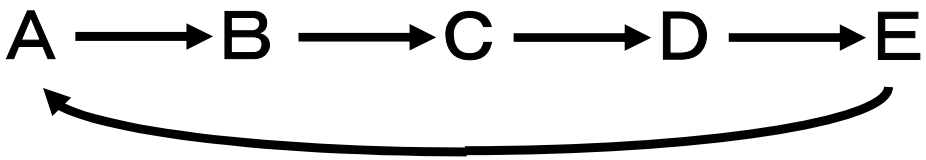
☆Cascade leads to multiplicative propagation of noise (at critical region)



d Nx/dt=Ny N **z**

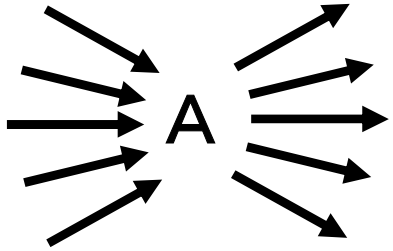
with cascade catalytic reactions, fluctuations are successively multiplied,
(cf addition in central limit theorem.);
Hence after logarithm, central limit th. applied

☆Cascade leads to multiplicative propagation of noise (at critical region)



Propagation of fluctuation, feedback to itself, leading to log-normal distribution tail.

Cf. If parallel,



Cf??

weight – log-normal
height -- normal

Fluctuations come in parallel:

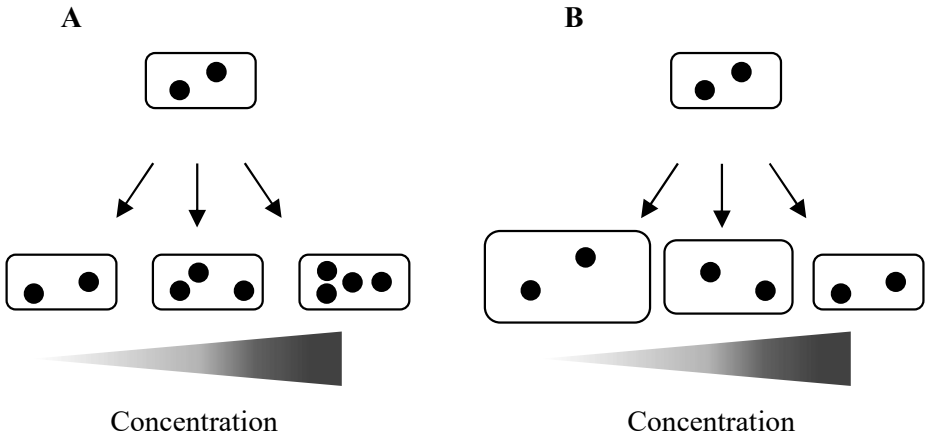
Usual central limit theorem is valid;

normal distribution.

Growth Fluctuation induces log-normal-type distrb.

Figure 1

Fluctuations in a Cell; Cell Volume Growth effect



Stochastic gene expression that are current concern of many

Consequence of Cell volume growth fluctuation tha we are interested

Tsuru,Ichinose,Kashiwagi,Ying,KK,Yomo

Origin of Log-tailed phenotypic fluctuation

- protein concentration x
- dx/dt=f(x)-(μ+η）x

dilution term by cell volume growth

μ ——growth rate

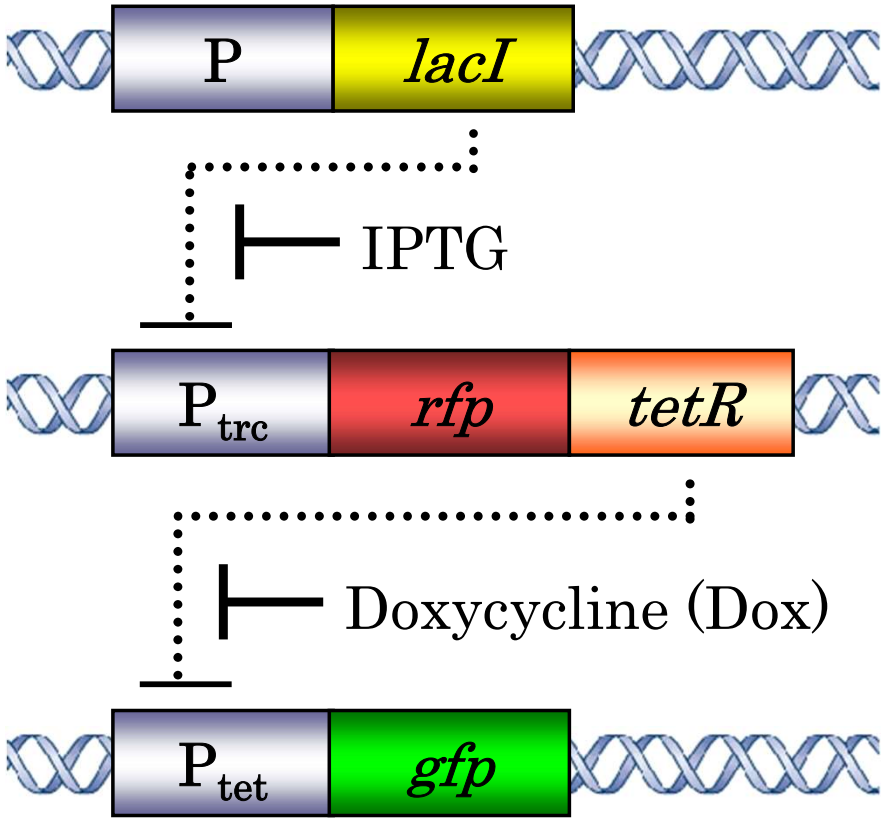
η ——fluctuation (noise)

multiplicative noise → log-tailed distribution

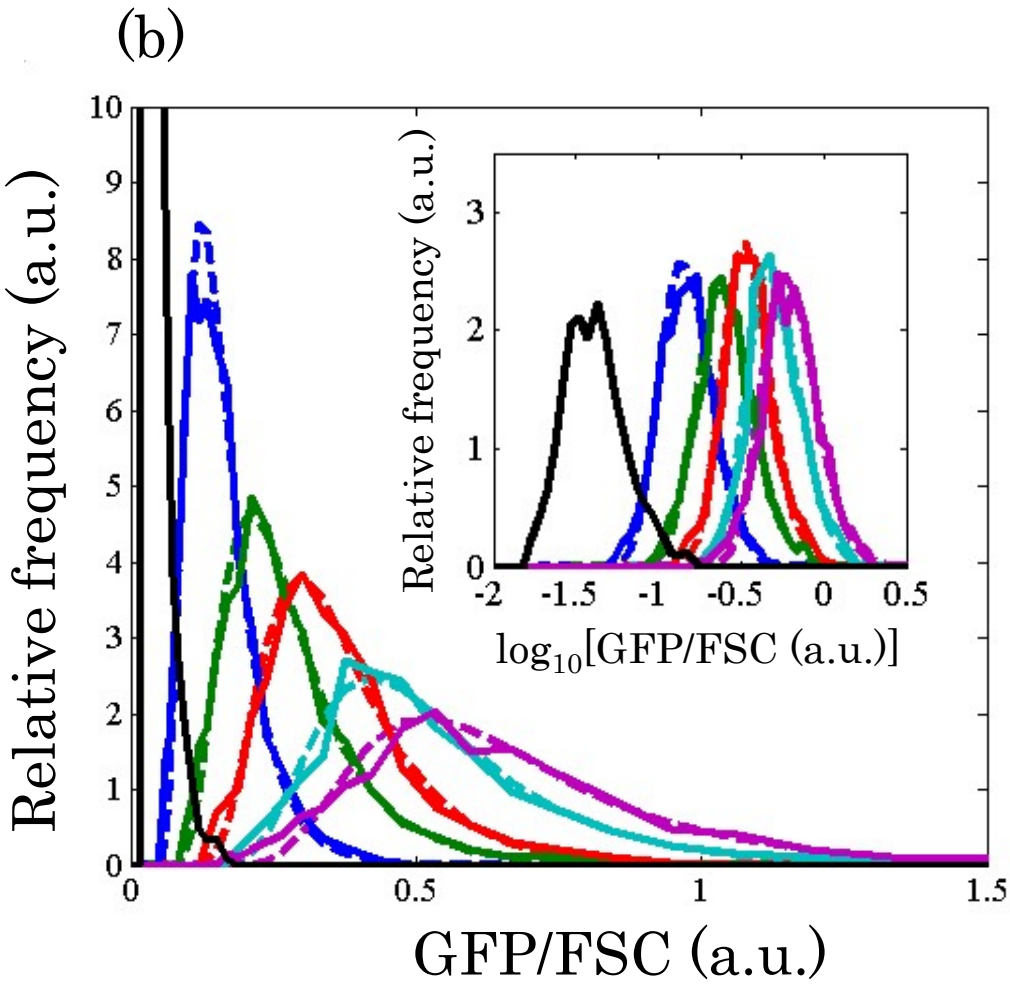
(exp； Tsuru etal）

Growth rate μ is a result of an ensemble of gene expression μ(x1,x2,x3,…) --(consistency)?

Statistics in gene expression in the present cell



Chromosome in *E. coli*



Log-normal like distribution at each Doxycycline concentration

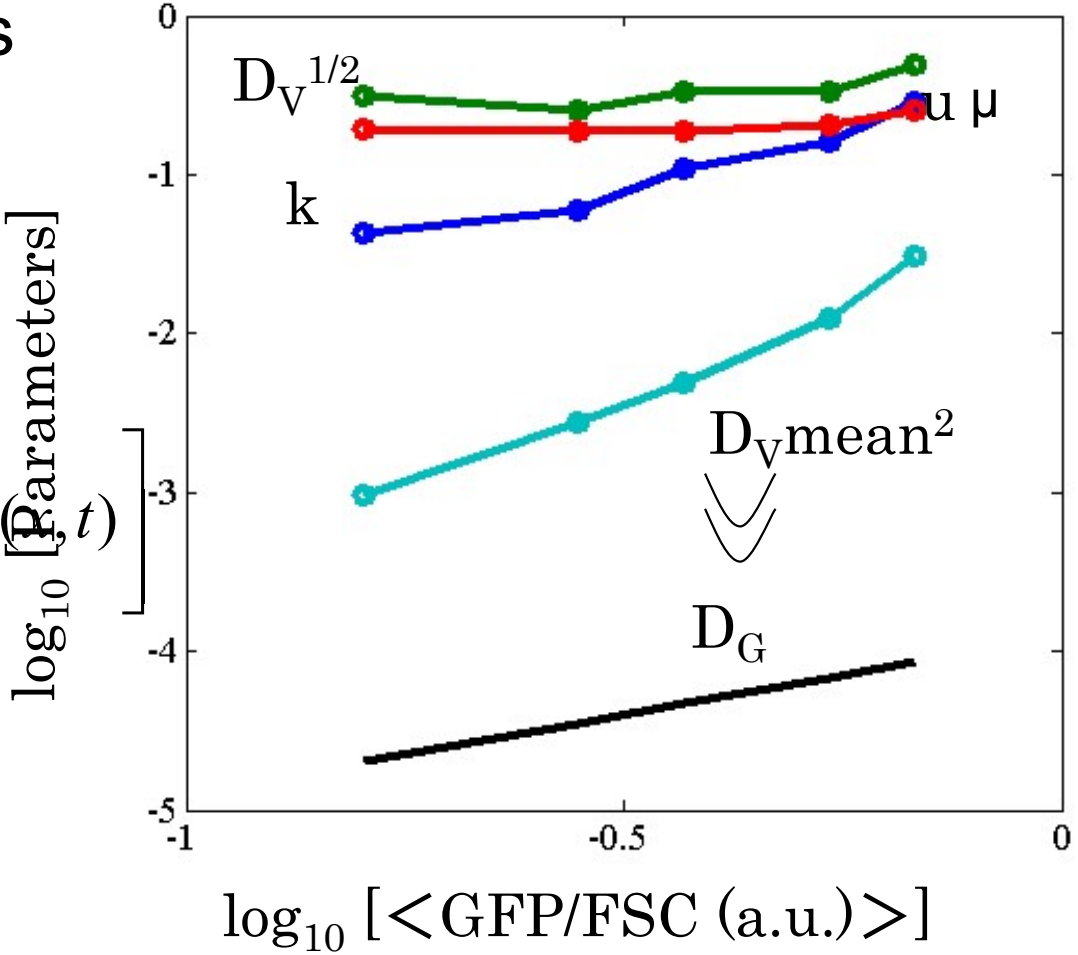
Tsuru,Ichinose,Kashiwagi,KK, Yomo

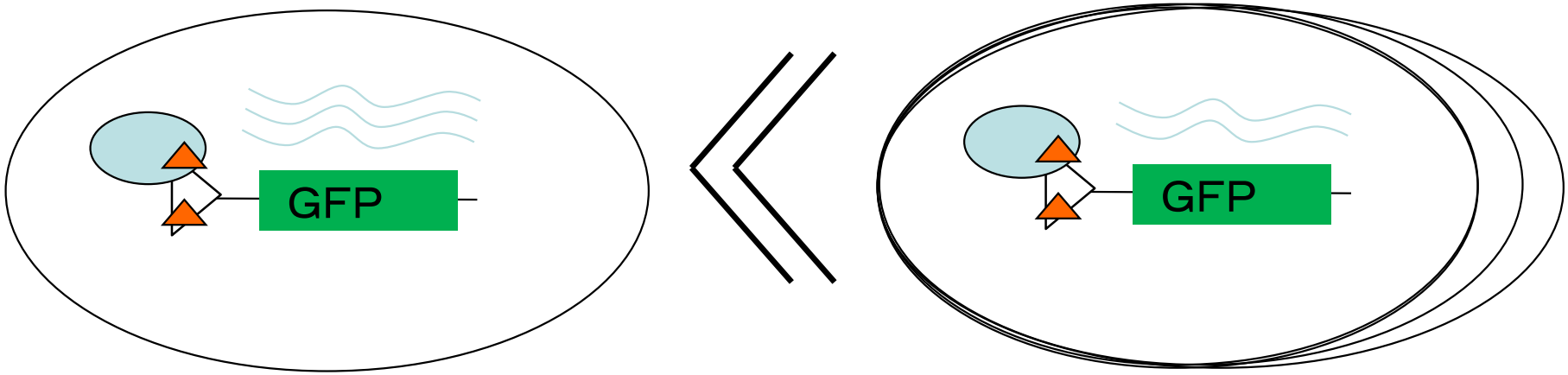
Simple interpretation by growth-rate
Fluctuation → dilution fluctuates

$$\frac{dx}{dt} = k - \left(\mu + \eta_V(t)\right)x + \eta_G(t)$$

$$\left\langle \eta_i(t)\eta_i(t')\right\rangle = 2D_i\delta(t-t')$$

$$\frac{\partial P(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[\left(\mu - D_V\right)\left(x - \frac{k}{\mu - D_V}\right)P(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[\left(D_G + D_V x^2\right)P(x,t)\right]$$





$$\frac{\partial P(x,t)}{\partial t} = \frac{\partial}{\partial x}\left[\left(\mu - D_V\right)\left(x - \frac{k}{\mu - D_V}\right)P(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[\left(D_G + D_V x^2\right)P(x,t)\right]$$

(3)

Temporal changes in statistical moments are calculated as follows:

$$\frac{d\langle x\rangle}{\partial t} = -\left(\mu - D_V\right)\left(\langle x\rangle - \frac{k}{\mu - D_V}\right)$$

(4.1)

$$\frac{d\langle x^2\rangle}{\partial t} = -2\left(\mu - 2D_V\right)\langle x^2\rangle + 2\left(k\langle x\rangle + D_G\right)$$

(4.2)

, where

$$\langle x\rangle = \int_{-\infty}^{\infty} xP(x,t)dx$$

and

$$\langle x^2\rangle = \int_{-\infty}^{\infty} x^2 P(x,t)dx$$

.

The solutions at steady state are analytically solved as follows.

$$\langle x\rangle_{st} = \frac{k}{\mu - D_V} \quad (5.1)$$

$$P(x) = \frac{C}{x^{1+\frac{\mu}{D_V}}}\exp\left[-\frac{k}{xD_V}\right]$$

Relaxation rate → μ—Dv

Average x → k/(μ—Dv)

Variance Dv,Dg separetdely

$$\left\langle x^2 \right\rangle_{st} = \frac{k\left\langle x \right\rangle_{st} + D_G}{\mu - 2D_V}$$

(5.2)

The temporal solution of average is solved as follows.

$$\left\langle x \right\rangle = \left\langle x \right\rangle_{st} - \left(\left\langle x \right\rangle_{st} - \left\langle x \right\rangle_0 \right) \exp\left[-\left(\mu - D_V \right) t \right]$$

represents the average at time zero. The relative fluctuation, CV^2, which is defined as the variance divided by square of average, at the steady state is also solved as follows. $\left\langle x \right\rangle_0$

$$CV^2 = \frac{\left\langle x^2 \right\rangle_{st} - \left\langle x \right\rangle_{st}^2}{\left\langle x \right\rangle_{st}^2} = \frac{1}{\mu - 2D_V} \left(D_V + \frac{D_G}{\left\langle x \right\rangle_{st}^2} \right)$$

$$CV^2 \approx \frac{D_V}{\mu - 2D_V} \qquad \left(\left\langle x \right\rangle_{st}^2 D_V \rangle\rangle D_G \right)$$

But CV2 is almost constant then contribution by Dg is small

Growth rate μ does not change so much by Doxy conc.

From the distribution, and temporal course estimate k,Dv,Dg

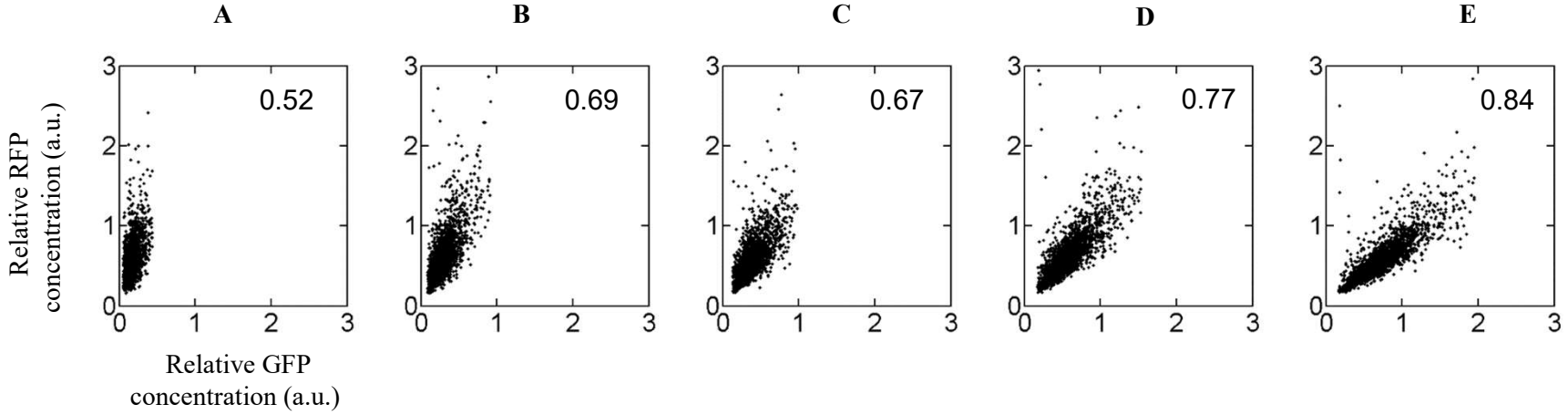
P(x) is also fitted well Dg is also estimated by stochastic gene expression analysis by Poissonian molecular process

Growth-rate gives global noise

RFP-GFP concentration correlation

(Just from gene network, negative correlation is expected, but clearly positive correlation is observed)

in the presence of doxycycline of various concentrations, 16.7 nM (A) , 22.5 nM (B), 33.7 nM (C), 45 nM (D) and 113 nM (E).



$$C_{xy} = \frac{\langle x^2 \rangle - \langle x \rangle \langle y \rangle}{\sqrt{\langle x^2 \rangle - \langle x \rangle^2}\sqrt{\langle y^2 \rangle - \langle y \rangle^2}}$$

Negative Feedback: higher growth --- higher dilution for all proteins --- lower cellular activity ---lower growth

Large fluctuation in growth rate (30-50%)

Question:: Source for growth fluctuation?

Furthermore, Time Scale for the growth fluctuation is rather slow (far from white noise, order of cell division):

(the stationary distribution is not much affected)

(Ichinose et al)

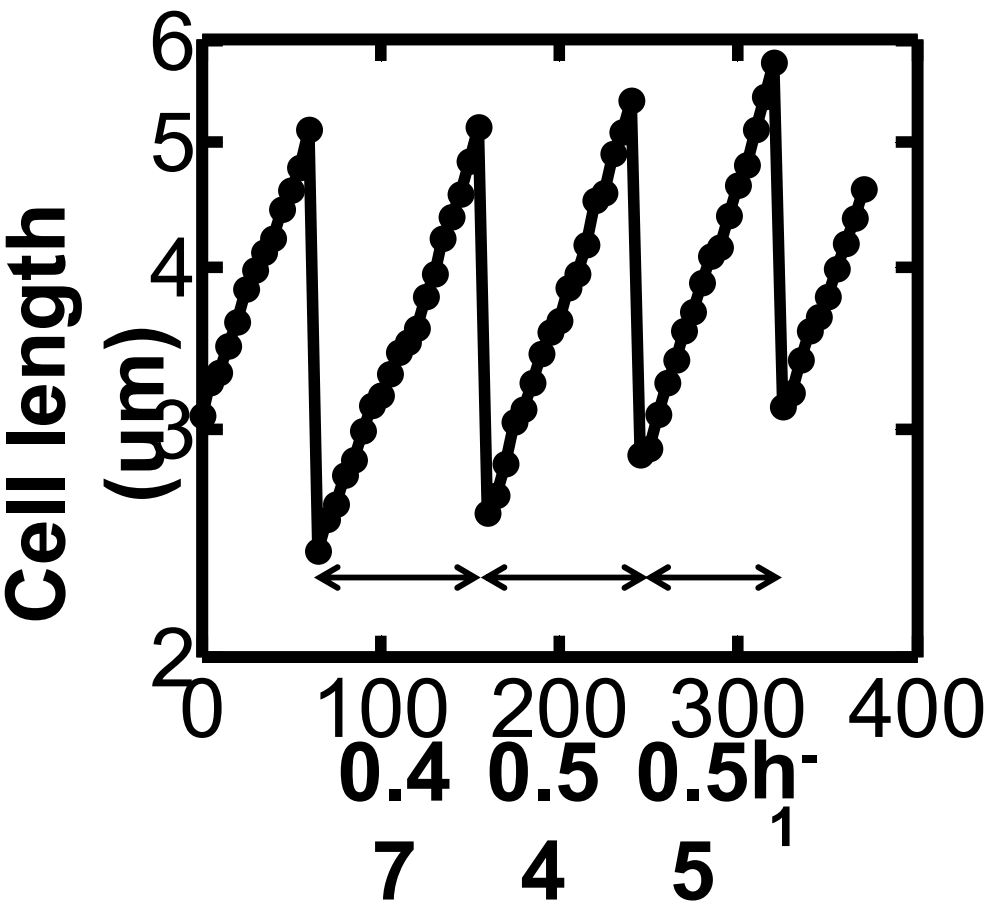
$$\frac{dx}{dt} = k - \mu x \qquad x = \frac{k}{\mu} \qquad dx = -\frac{k}{\mu^2} d\mu$$

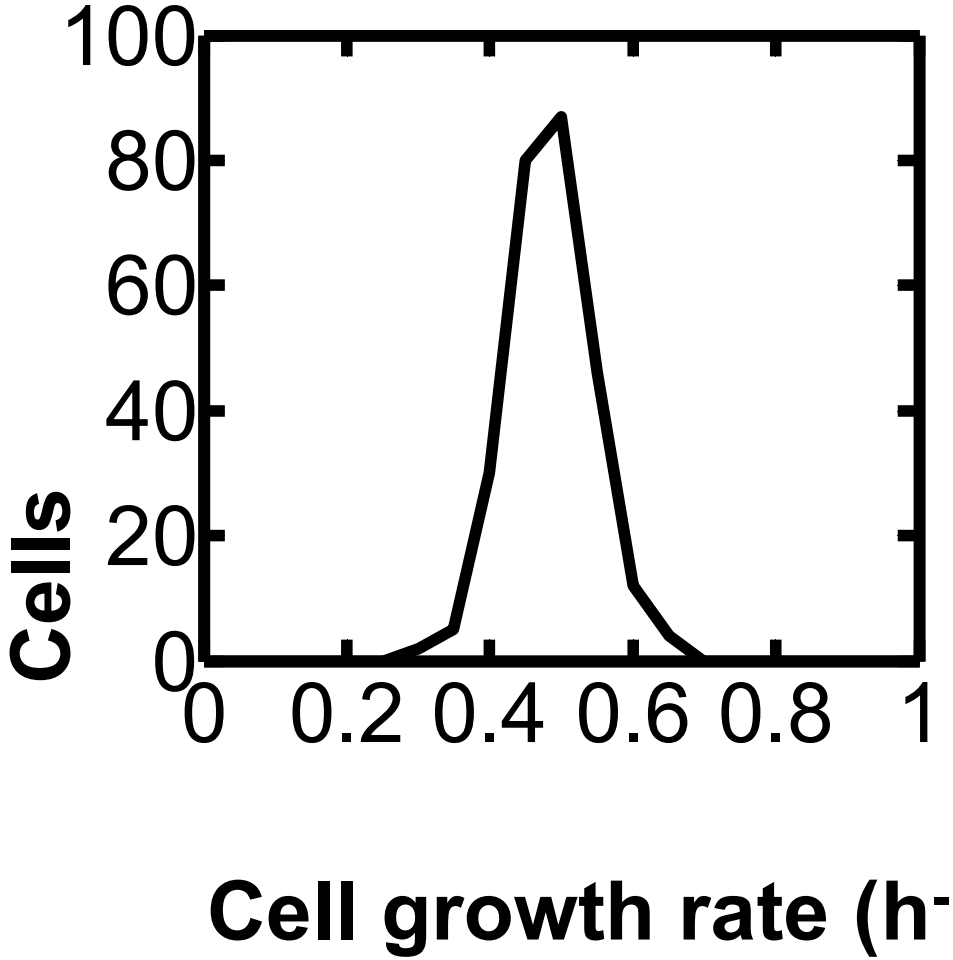
$$P(x)dx = P(\mu)d\mu$$

$$P(x) = P(\mu)\frac{d\mu}{dx} = \left|-\frac{\mu^2}{k}\right| \frac{1}{\sqrt{2\pi\sigma_\mu{}^2}} \exp\left[-\frac{(\mu-\mu_0)^2}{2\sigma_\mu{}^2}\right]$$

$$= \frac{k}{x^2} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{\left(\frac{k}{x}-\mu_0\right)^2}{2\sigma_\mu{}^2}\right]$$

Fluctuation in growth rates





CV=13%

Correlation time ～1 cell generation (nonwhite)