Reproduction: consistency between molecule replication and cell reproduction

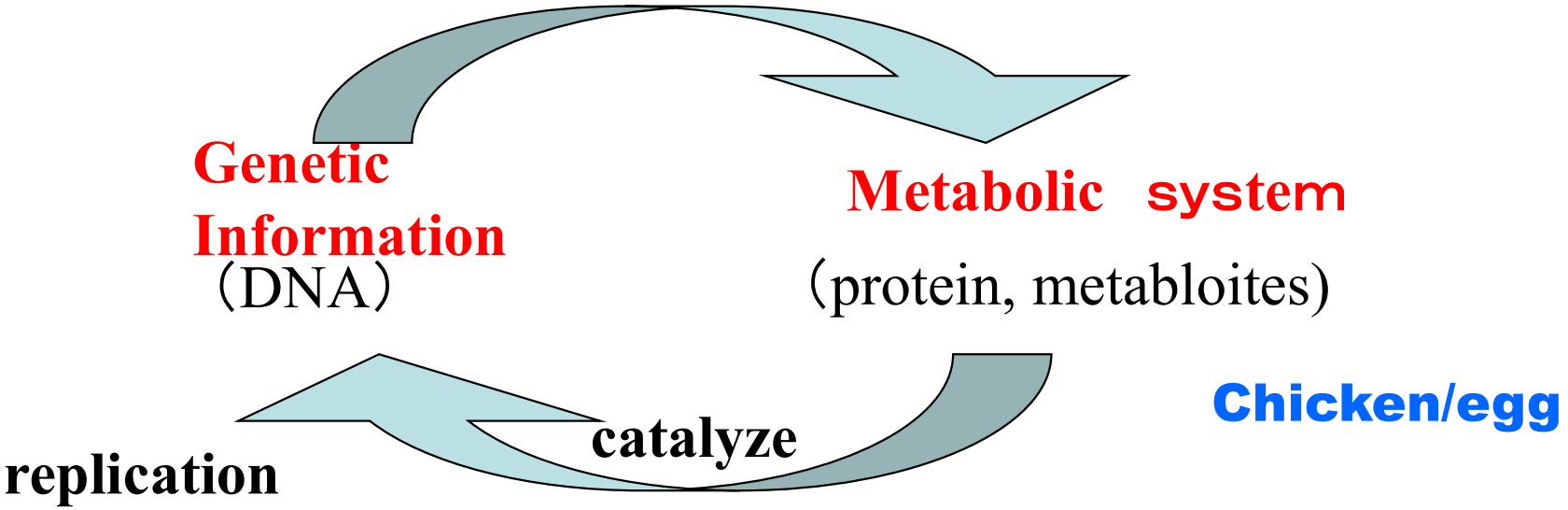
(Probably I'll skip last few slides)

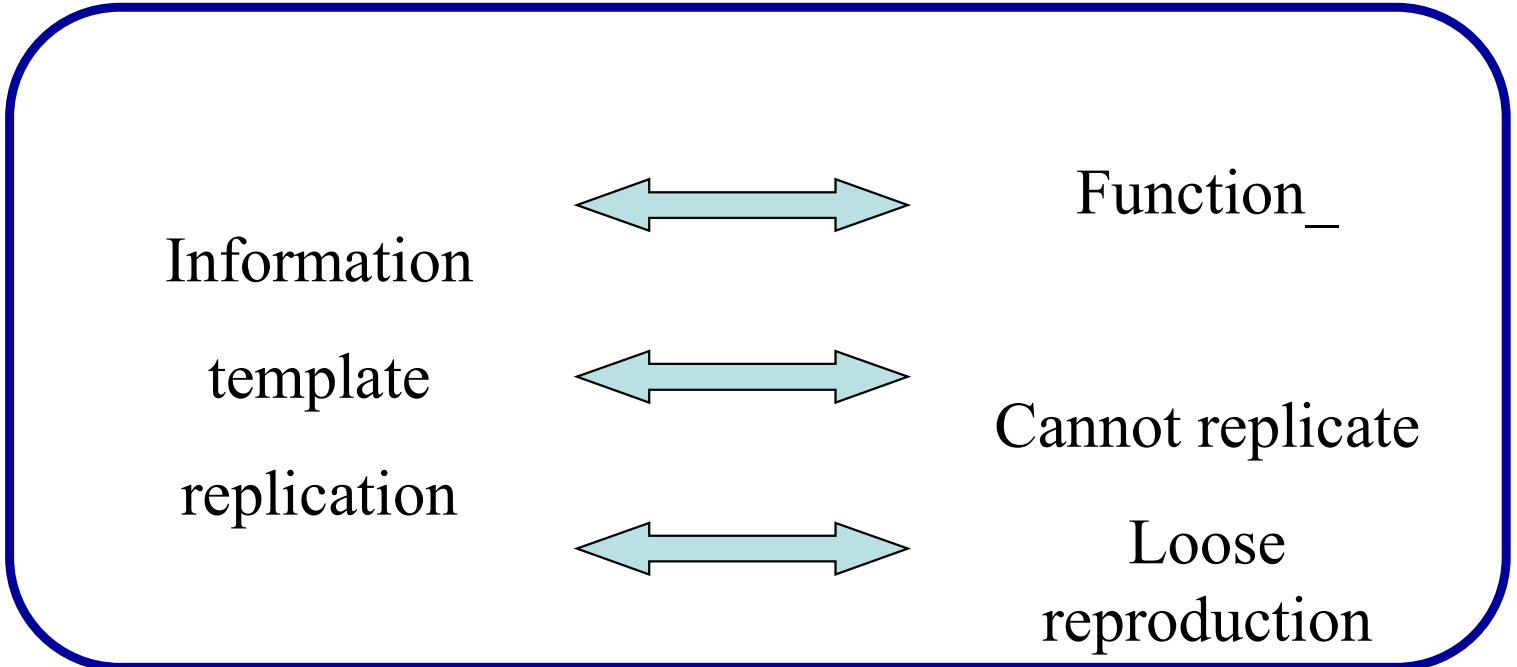
- Construction of Reproducing Cells/ Origin of Life
- Nonequilibrium?—encapsulated enzyme
- Some macroscopic laws

- Dormant state as breakdown of consistency

- Constructive biology – primitive cells
  origin of central dogma, origin of compartment
  diversity

Genetic Information vs Metabolism first ?





- Replication 1st vs Metabolism 1st ?

- Replication of Information Molecule vs Maintenance of Complex System

  (Manfred Eigen vs Freeman Dyson)

chemist physicist

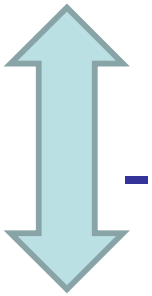
From Loose Reproduction to Genetic Takeover?

Instead of answering which is first, but discuss how the two are compatible, how diversity is sustained and how symmetry breaking between information and function (catalysis) has arisen

- Reproduction of cells with Divers Components

- Steps from a set of catalytic reactions to protocells

Just a set of catalytic reactions

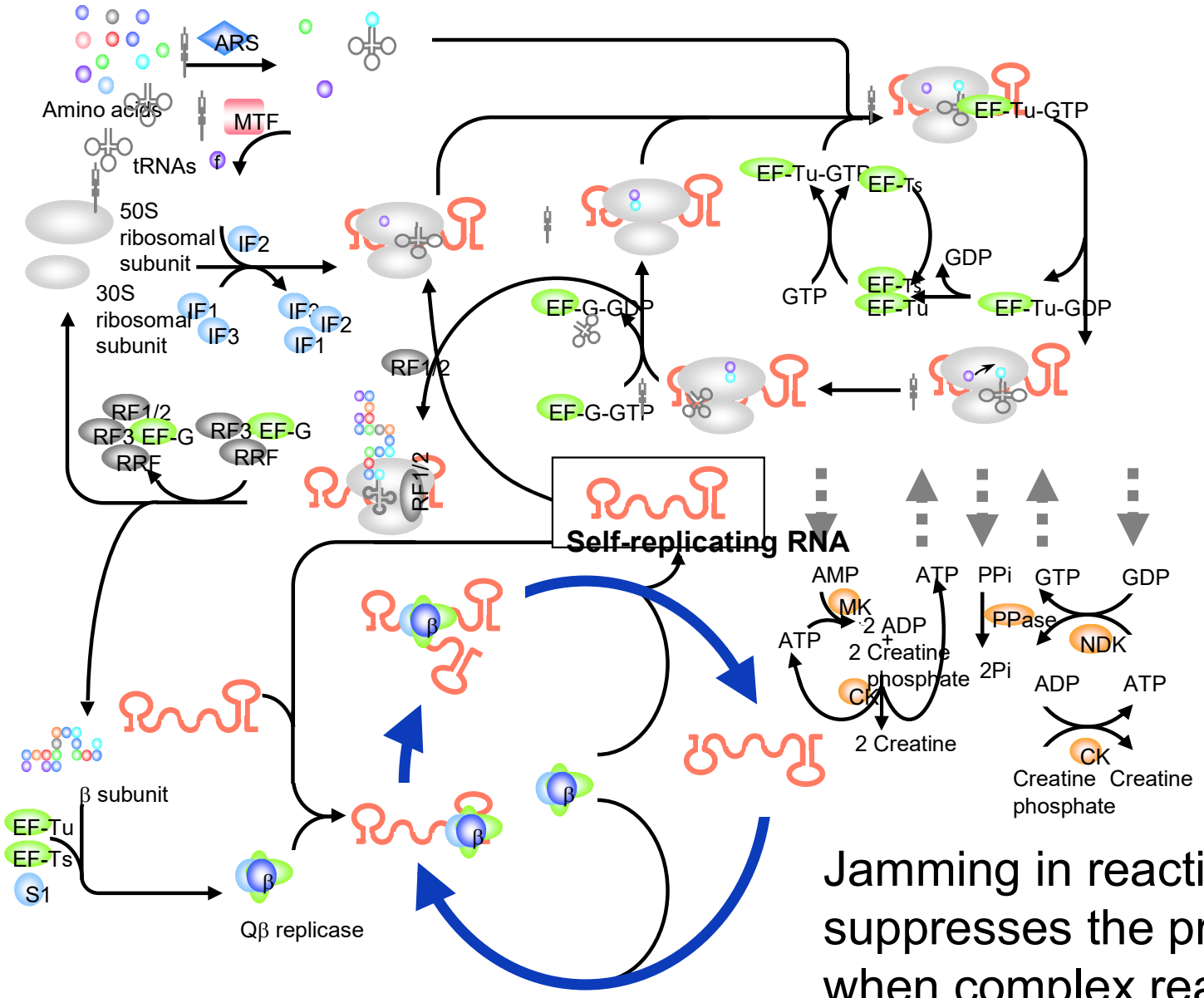


----- Still, a large gap ----

Cell (autonomously reproduces itself, adapts to new environment, evolves, ….)

??Fill the gap both in theory and experiments

E.g. Pure system: More than 5000 reaction steps run with 144 species of bio-molecules for the in vitro self-replication (extracted from E coli)



Pure System

Jamming in reaction by crowding often suppresses the process, especially when complex reaction is put

- Still, many more steps in experiments:

Theoretical issue?

Life needs very fine-tuning, or does it belong to a type of 'universality class'? →

- We need Basic Theoretical Concepts to solve basic questions

Basic concepts already proposed :

error catastrophe/threshold, hypercycle, quasispecies (Eigen)

loose reproduction, autocatalytic-set(Dyson,Kauffman)

Need more?

Minority-control, chemical-net glass, consistency between cell growth and molecular replication, discreteness-induced transition,

IdealCellModel?

Compartment （Membrane）+

Metabolism（Catalysys） +

Information (Template)

- Nonequilibrium?
- Encapsulated enzyme (importance of compartment+ catalysis)
- Single outside bath in contrast to Carnot
- Minimum-model for replicating nonequlibrium system？
- Resource——＞Internal Catalysys-> Reproduction

- II: Timescale Problems: *or why growth favored*

Compartment (cell) catalyst only within

Most reactions are facilitated by catalysts

e.g., resource ←→ product (waste) possible

R+C<->P+C internal ○

external X

-R,P disconnected outside :

Any ratio is sustained (within normal time scale)

(Non-equilibrium, outside)

Cell - 'equilibration apparatus' to unveil external non-eqb condition by encapsulated enzyme?

- Enzyme as Timescale Controller

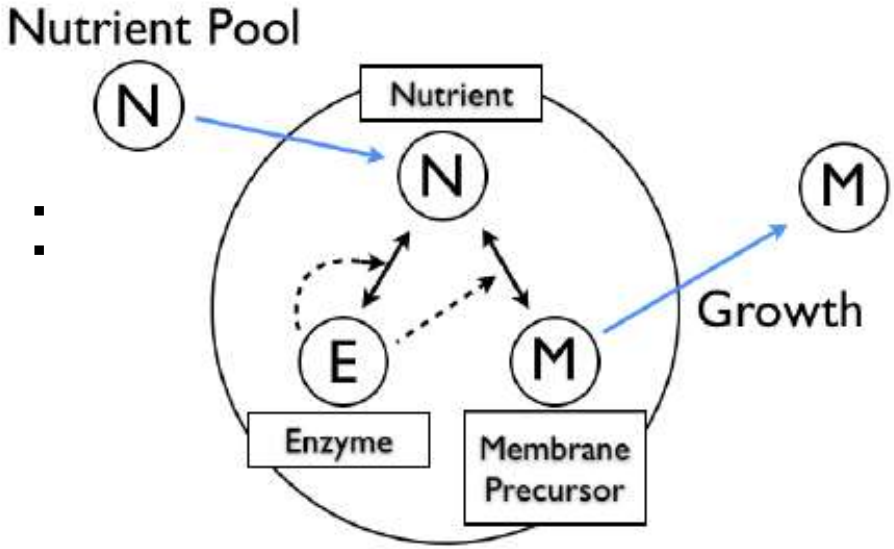
Compartment (cell) catalyst only within

Most reactions are facilitated by catalysts

Catalysts often assumed that does not change equilibrium condition, but when time scale is reduced 10^{12} or so, then effectively

$N+C \rightleftharpoons P+C$ internal O

external X



-N,P disconnected outside :

Any ratio is sustained

(within normal time scale)

Non-equilibrium, outside

$$\dot{E} = E(R - k_E E) - \mu E$$

$$\dot{L} = cE(R - k_L E) - \gamma L - \mu L$$

$$\dot{M} = \gamma L - \mu M$$

$$R + E \rightleftharpoons E + E$$

$$R + E \rightleftharpoons L + E$$

$$\mu \sim \frac{\dot{V}}{V} \rightarrow \gamma L$$

$\downarrow$

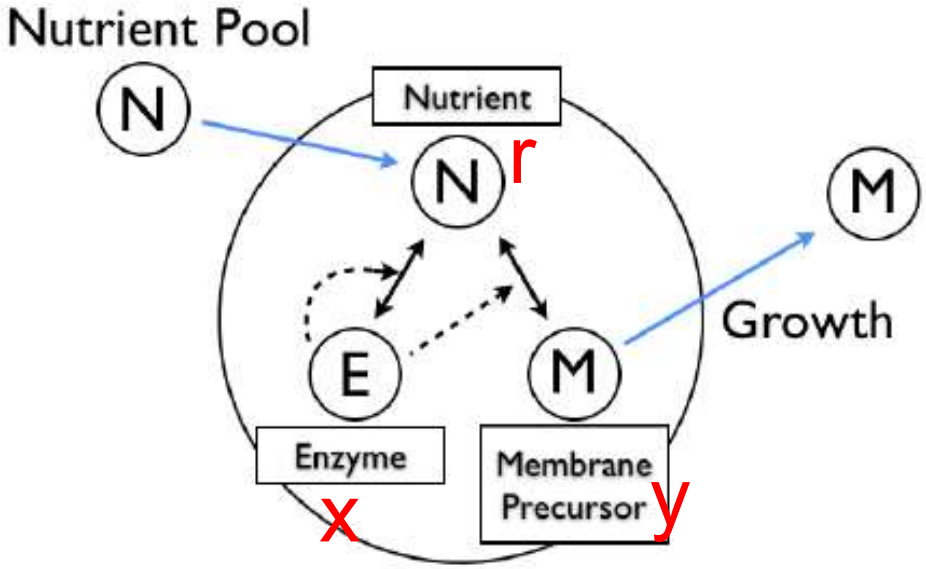
M = membrane → cell volume

L/E balanced growth

Growth-rate μ √R not agree with Monod's law

A toy model with resource + enzyme+ membrane

（Himeoka,KK 2014、PRE）



For higher enzyme abundances
→ Approach equilibrium → smaller entropy production
Higher Flow -> Lower Loss?

x,y concentrations of enzyme and membrane precursor

E⇆ N dilution

$$\frac{dx}{dt} = \kappa_x x(kr - x) - x\lambda,$$

$$\lambda \equiv \frac{1}{V}\frac{dV}{dt}$$ Growth-rate

$$\frac{dy}{dt} = \kappa_y x(lr - y) - \phi y - y\lambda.$$

N⇆Mp Mp →M dilution

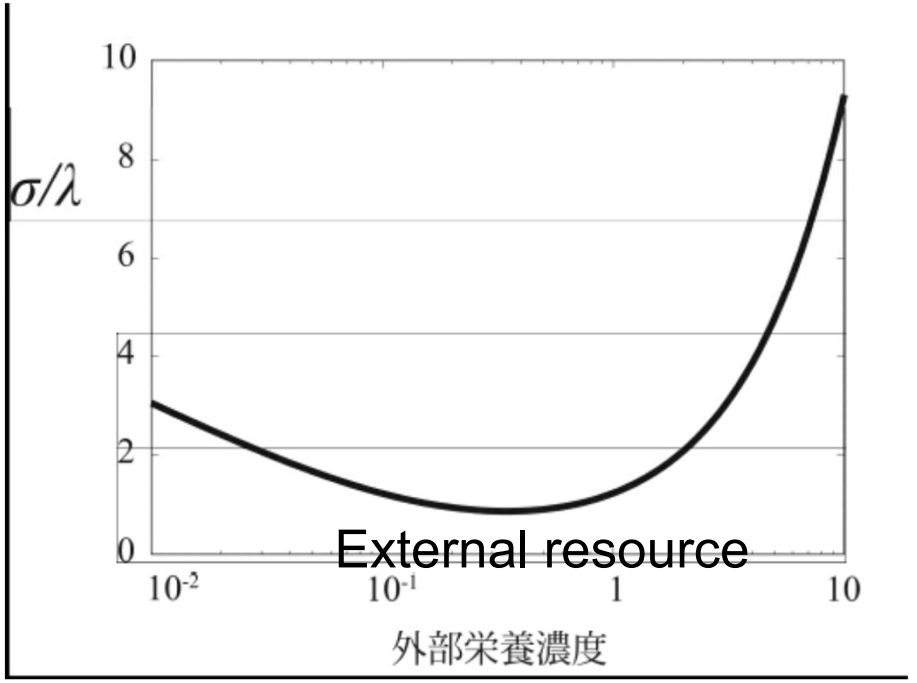
$$\frac{dx}{dt} = \kappa_x x(kr - x) - x\lambda,$$

$$\frac{dy}{dt} = \kappa_y x(lr - y) - \phi y - y\lambda.$$

$$\lambda \equiv \frac{1}{V}\frac{dV}{dt}$$

component i $(i = x, y)$ [33]; and ϕ is the consumption rate of the membrane precursor to produce the membrane such that the volume growth rate λ is given by $\lambda = \gamma\phi y$, where γ is the conversion rate from membrane molecules to the cell volume. by $\sigma = \sum_i J_i \frac{A_i}{T}$, where J_i is the chemical flow and A_i is the affinity for each reaction. Here we set $T = 1$ without loss of

$$\tilde{\sigma} = \tilde{\kappa}\tilde{x}(\tilde{k}\tilde{r} - \tilde{x})\ln(\tilde{k}\tilde{r}/\tilde{x}) + \tilde{\kappa}\tilde{x}(\tilde{r} - \tilde{y})\ln(\tilde{r}/\tilde{y}).$$

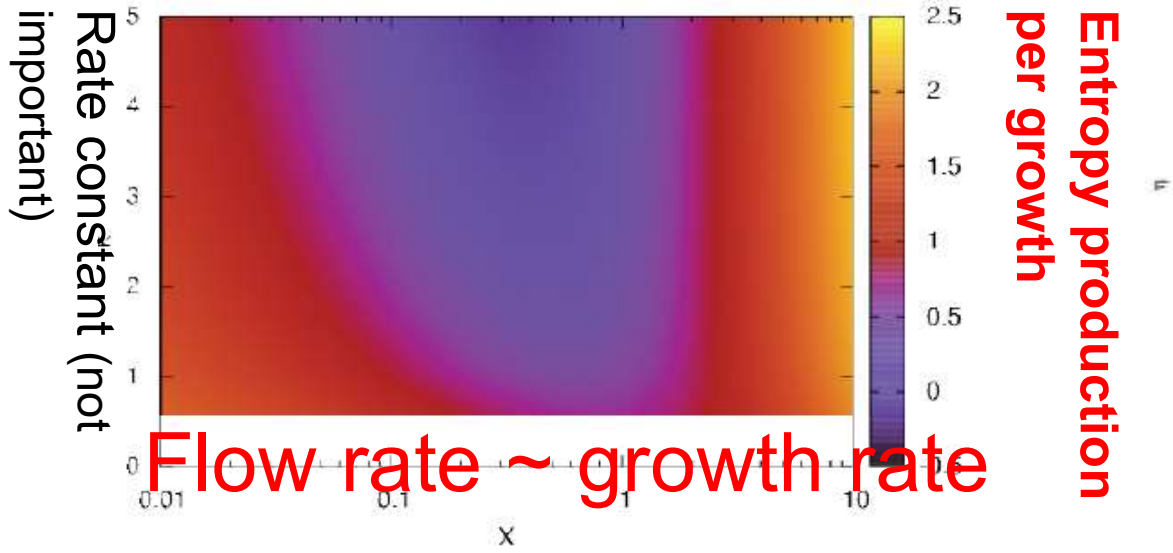


Thermodynamic loss is minimum at a finite flow, i.e. with cell growth

- Compute entropy production

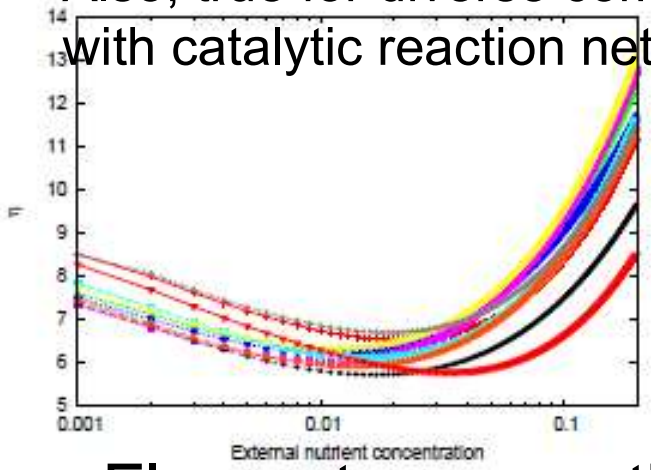
by $\sigma = \sum_i J_i \frac{A_i}{T}$, where J_i is the chemical flow and A_i is the affinity for each reaction. Here we set $T = 1$ without loss of

$$\tilde{\sigma} = \tilde{\kappa}\tilde{x}(\tilde{k}\tilde{r} - \tilde{x})\ln(\tilde{k}\tilde{r}/\tilde{x}) + \tilde{\kappa}\tilde{x}(\tilde{r} - \tilde{y})\ln(\tilde{r}/\tilde{y})$$



Flow rate ~ growth rate

Also, true for diverse components with catalytic reaction network

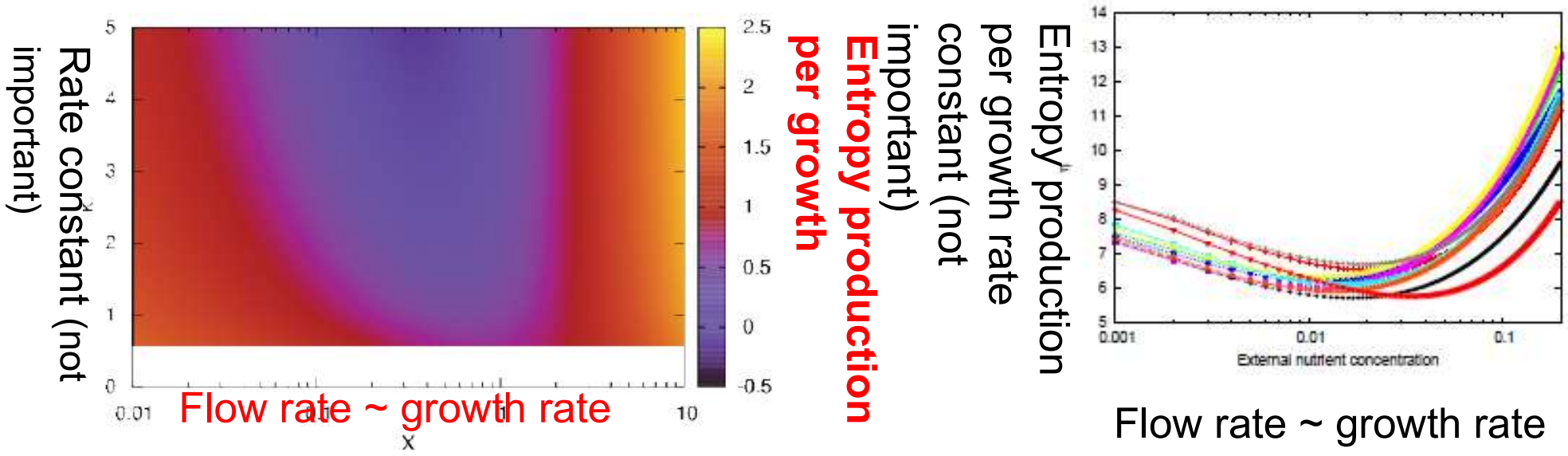


Flow rate ~ growth rate

Why?: higher flow → more enzyme→enhances equilibration → decrease loss

Much higher: growth-dilution → loss increases by flow

Growth+ Dilution → Minimal Loss at Optimal flow rate



Autonomous regulation of time scale by enzyme concentration +
Loss by Growth

So far energetically not as a mass

Cell - ‘equilibration apparatus’ to unveil external non-eqb condition by encapsulated enzyme?

→Cell Machine is totally different from Carnot-type engine (in which quasi-static process is most efficient)

Consider Minimum setup +

Thermodynamic Consequence of such machine

Laws in reproducing cells?

Micro-Macro consistency->laws

Growth Rate= Macro-Order Parameter

——> Dilution –'mean-field'

1）Laws in Growth Rates

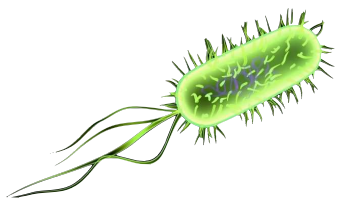
Monod,Pirt,Schaechter-Scott-Hwa

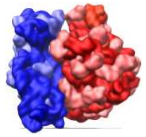
2） Wastes inevitable

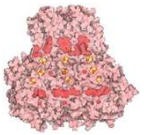
3）Efficiency at Finite Flow rate(←→ Carnot cycle)

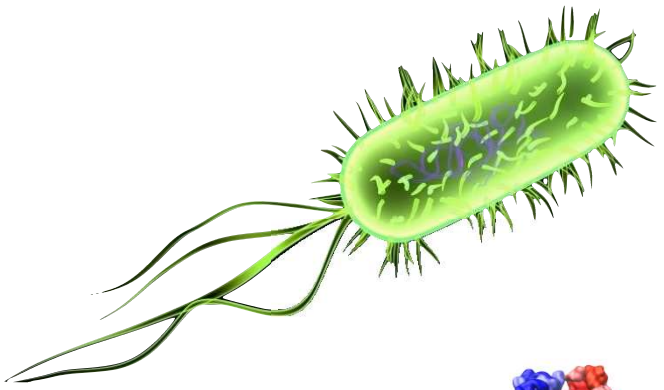
4） Exponential/Stationary/Dormant

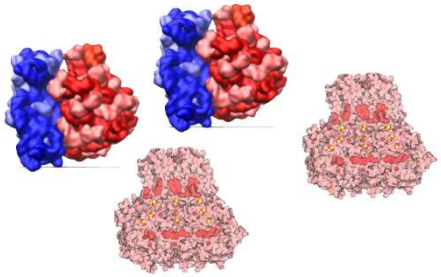
phase transition to 'sleeping states'



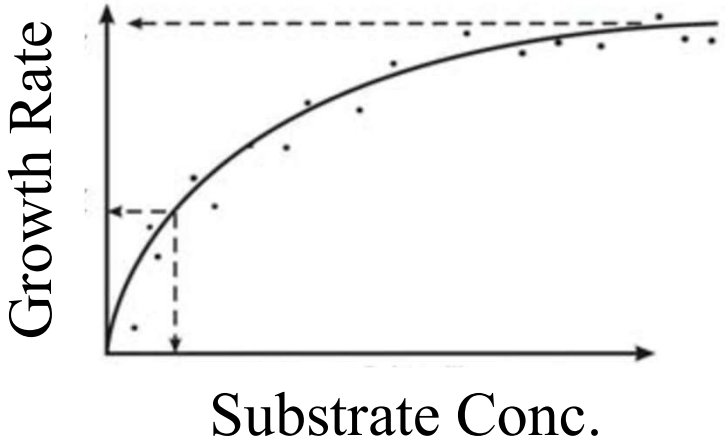




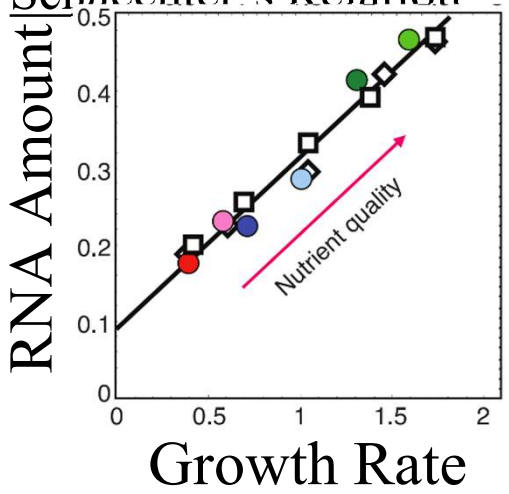




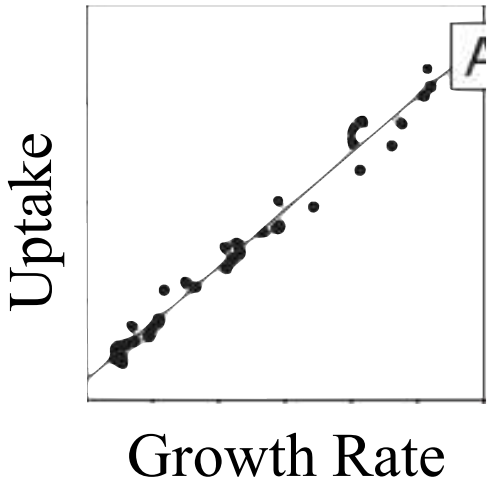
Monod's Growth law



Schaechter's Relation



Pirt's Equation



The Monod equation

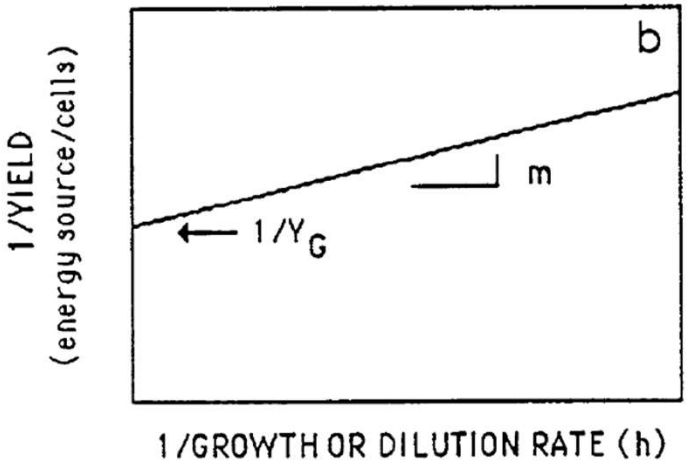
$$\mu = \mu_{max} S / (K_s + S)$$

Growth rate μ as a function of substrate concentration S.

μ_{max} K_s ： empirical coefficients depending on species and environmental conditions

cells/time). In Pirt's derivation, maintenance has no direct effect on growth rate, but the yield is decreased. On the basis of the assumption that Y is the actual yield of bacteria (grams of bacteria per gram of energy source) and Y_G is the theoretical maximum yield (Y if there were no maintenance), total energy utilization ($\mu x/Y$) can be partitioned into maintenance (mx) and true growth ($\mu x/Y_G$):

$$\mu x/Y = mx + \mu x/Y_G$$



and proposed a less hypothetical approach. The negative growth rate concept was circumvented by describing maintenance by a "coefficient" (m) that described the amount of energy needed to maintain cells for a given period (energy/cells/time). In Pirt's derivation, maintenance has no direct effect on growth rate, but the yield is decreased. On the basis of the assumption that Y is the actual yield of bacteria (grams of bacteria per gram of energy source) and Y_G is the theoretical maximum yield (Y if there were no maintenance), total energy utilization ($\mu x/Y$) can be partitioned into maintenance (mx) and true growth ($\mu x/Y_G$):

$$\mu x/Y = mx + \mu x/Y_G$$

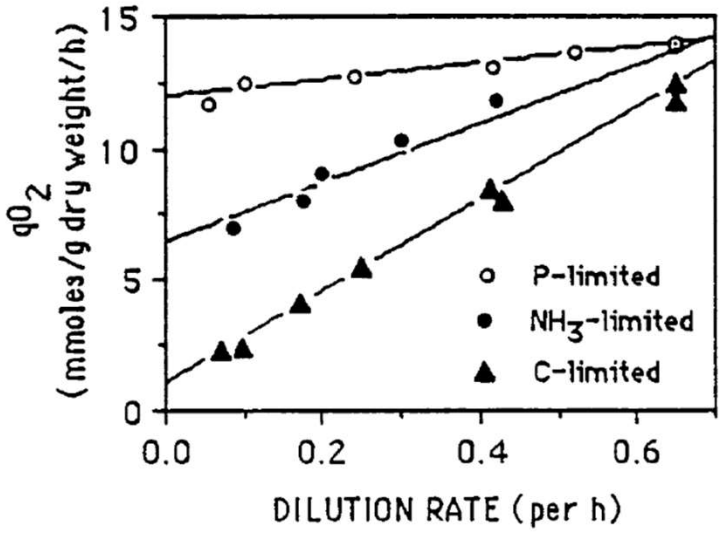


FIG. 4. Relationship between the specific growth rate and the specific rate of O_2 consumption in chemostat cultures of *Klebsiella aerogenes* growing in glucose-containing media that were limited by carbon, phosphorus, or ammonia. Reprinted with permission from reference 117.

Using the same type of algebraic transformations as Marr et al. (63), Pirt succeeded in deriving another straight-line equation (Fig. 3b):

$$\frac{1}{Y} = \frac{m}{\mu} + \frac{1}{Y_G} \quad (1)$$

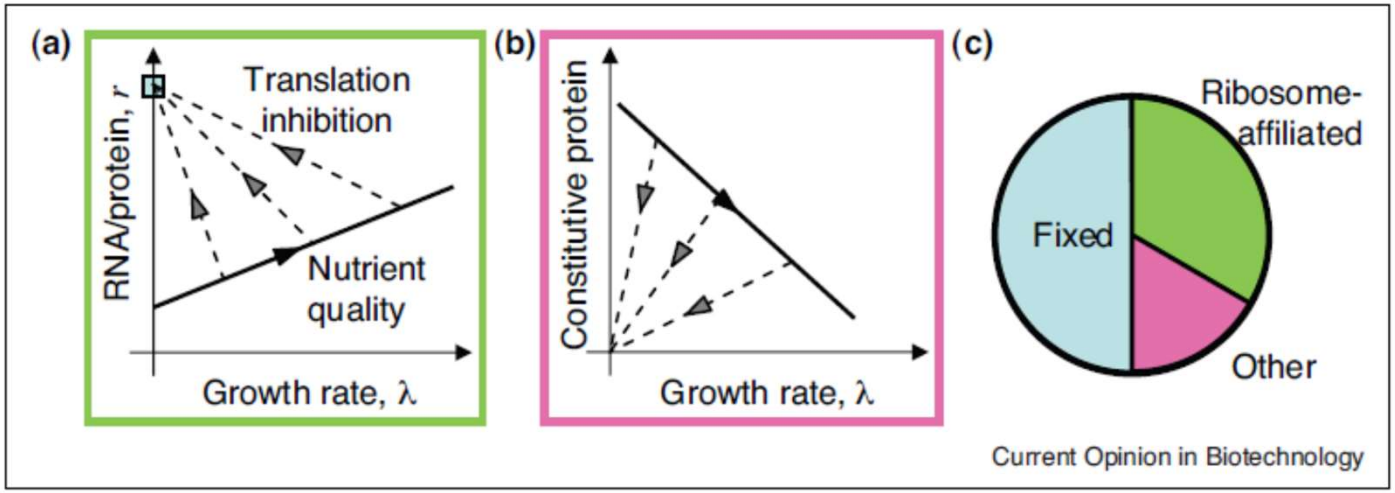
Resource->Growth (roughly linear relation)

Offset -- loss (cost for maintenace)

- RNA/Protein=aμ＋c

  for diverse conditions

  （Schaechter-Scott-Hwa (Curr Opinion.. 2011))



Ribosomal　Protein (self-replication）+ other prteins catalyzed by it

III: Issue of Waste disposal : (remark)

molecules without catalytic activities accumulate that may suppress others' catalytic activities (aggregation etc.)

→*Headache in artificial cell*

Growth-division process as garbage dump → growth/division necessary for maintenance?

Difficulty in reaching non-growth state without death? (*unless using 'glassy' relaxation*)

log→ stationary → dormant phases (death)

or differentiate by cell-cell interaction (symbiotic)

(partially supported by Toy Cell Model)

Problem of Waste Chemicals: growth/sleep

catalytic molecules are rare in polymer sequence

also mis-folding easily lead to loss of function

→ molecules without catalytic activities accumulate that may suppress others' catalytic activities (aggregation etc.)

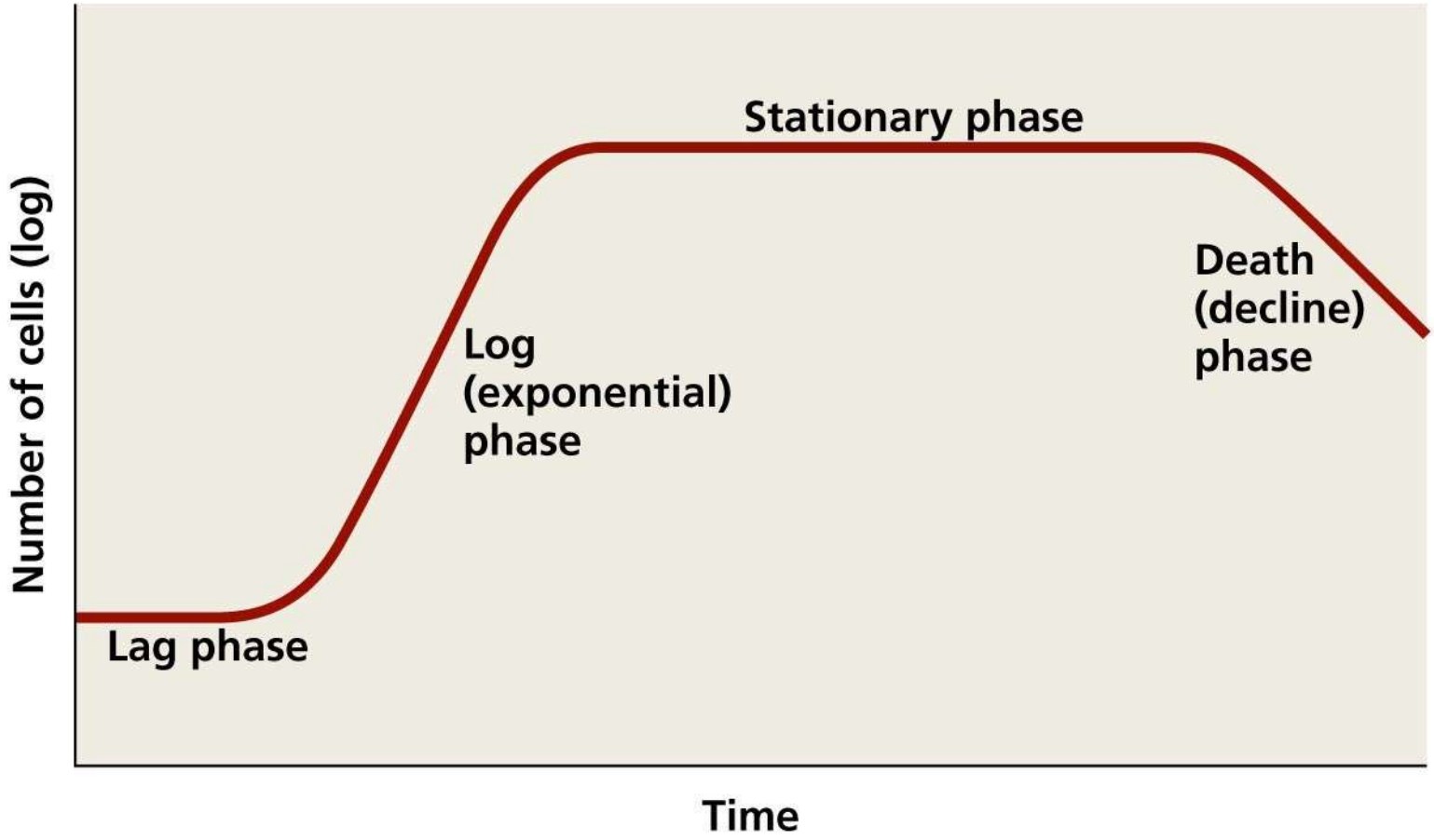
→Need Waste disposal

Growth-division process works as garbage dump

→ good for maintenance (*Growth hides every problem Churchill*)

…but without growth, waste dominates and dies?

Coexistence with (and regulation of) waste? → cells that can 'sleep'?


Number of cells (log)
Stationary phase
Death
(decline)
phase
Log
(exponential)
phase
Lag phase
Time

→Origin of 'Sleep' to go out of extinction

Question:

System with autocatalytic growth: dN/dt=aN

--- either growth (a>0) or death (a<0)

→Once resource is consumed, extinction follows

How to stop at a=0?? --- difficulty in keeping non-growth state without death. ..

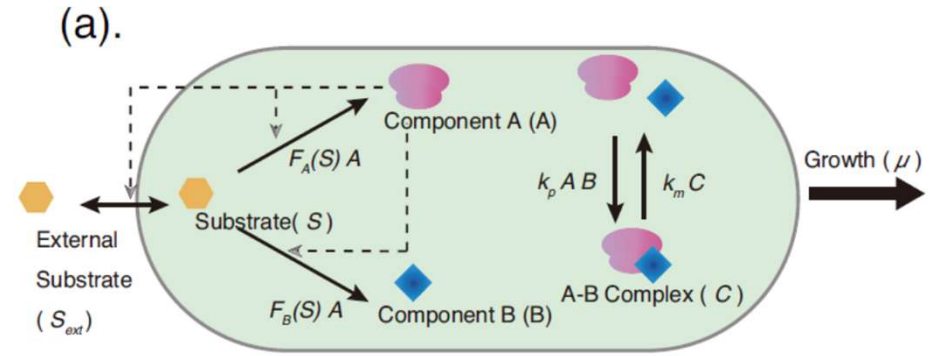
In present cells, transition from exponential-growth to stationary phases ('sleeping' state)

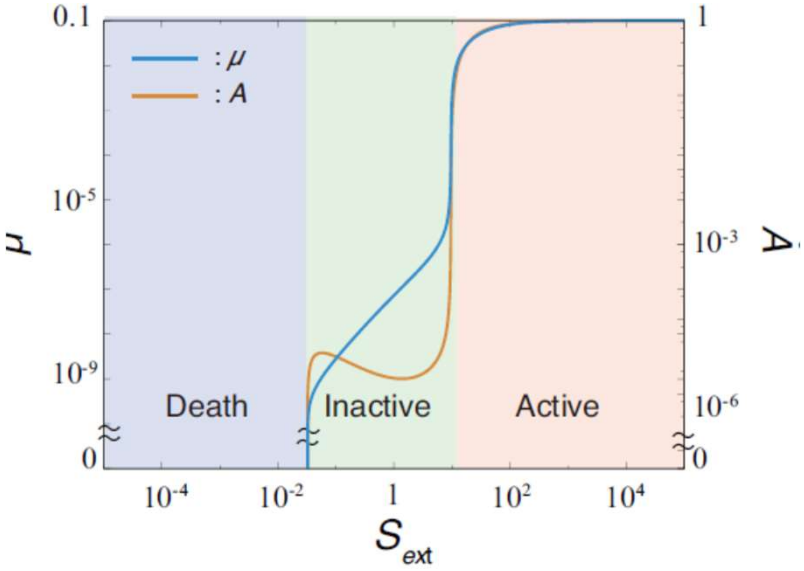
→inhibitory mechanism needed or Generic Mechanism? → regulation of time scale by 'waste' chemicals (Himeoka, KK, PhysRevX2018)

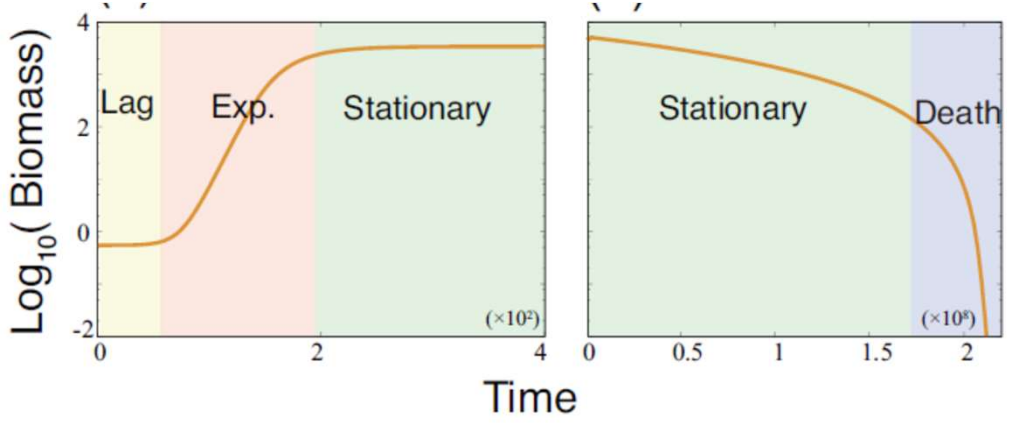
also relevant to artificial cells that can survive

Origin of sleep by Waste-Catalyst complex

Himeoka,KK2017







S: Substrate resource

A: active (autocatalytic) protein
eg ribosomal protein

B: Waste or inhibitor

S→ A, S→B catalyzed by P

A+B ⇄ Complex ('halting')

$$\frac{dS}{dt} = -F_A(S)A - F_B(S)A + A(S_{\mathrm{ext}} - S) - \mu S$$

$$\frac{dA}{dt} = F_A(S)A - G(A,B,C) - d_A A - \mu A$$

$$\frac{dB}{dt} = F_B(S)A - G(A,B,C) - d_B B - \mu B$$

$$\frac{dC}{dt} = G(A,B,C) - d_C C - \mu C$$

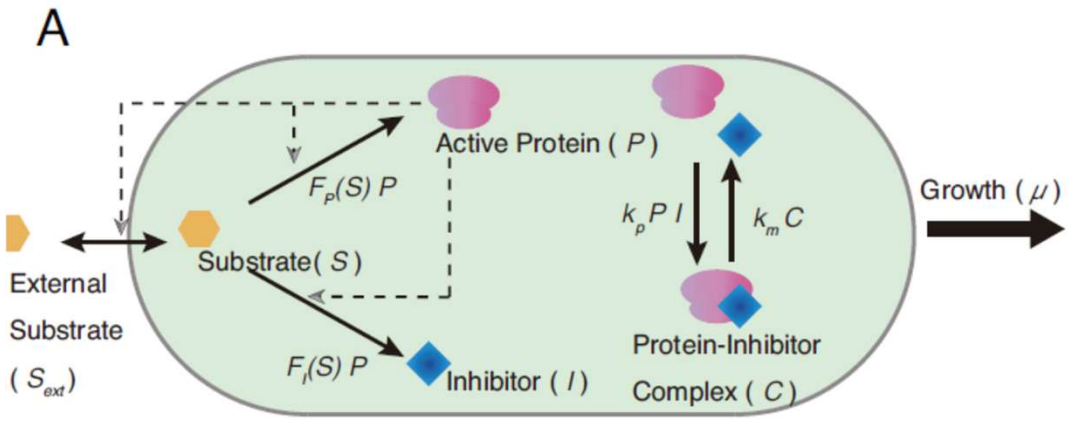
dC<dA,dB

Resource-rich → F(A)>F(B)
→autocatalytic by A

Limited → B accumulates to form complex C (halting)

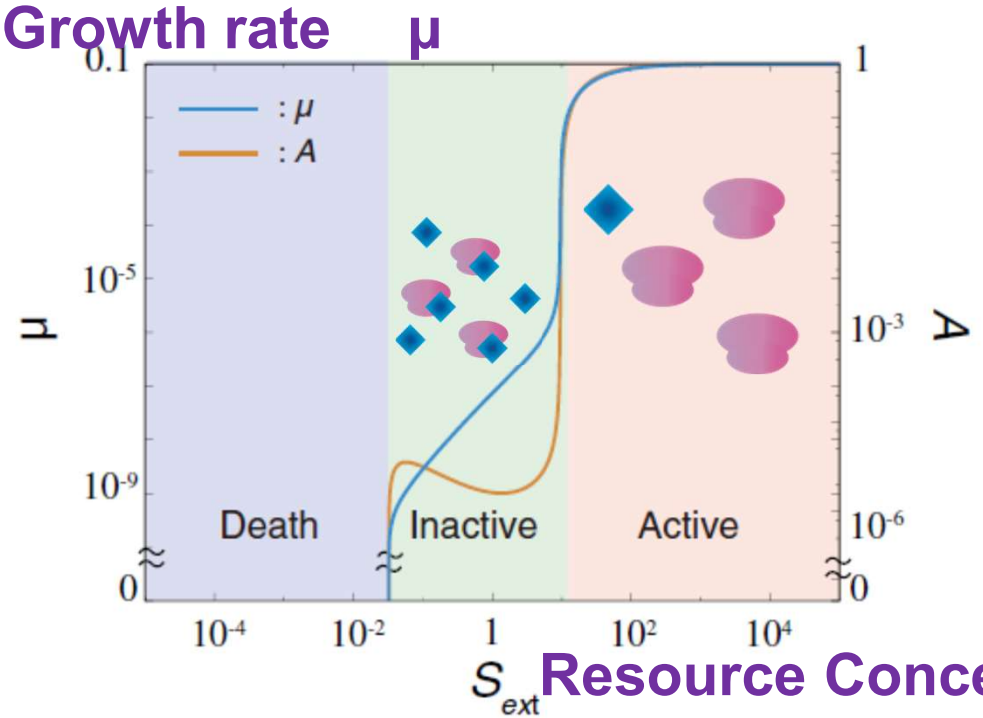
Activator for Growth +Inhibition by Waste → Transition to Sleeping state with Growth rate μ∼0 upon nutrient depletion

Himeoka,KK,2017 PRX



Most Active Proteins are trapped in Complex

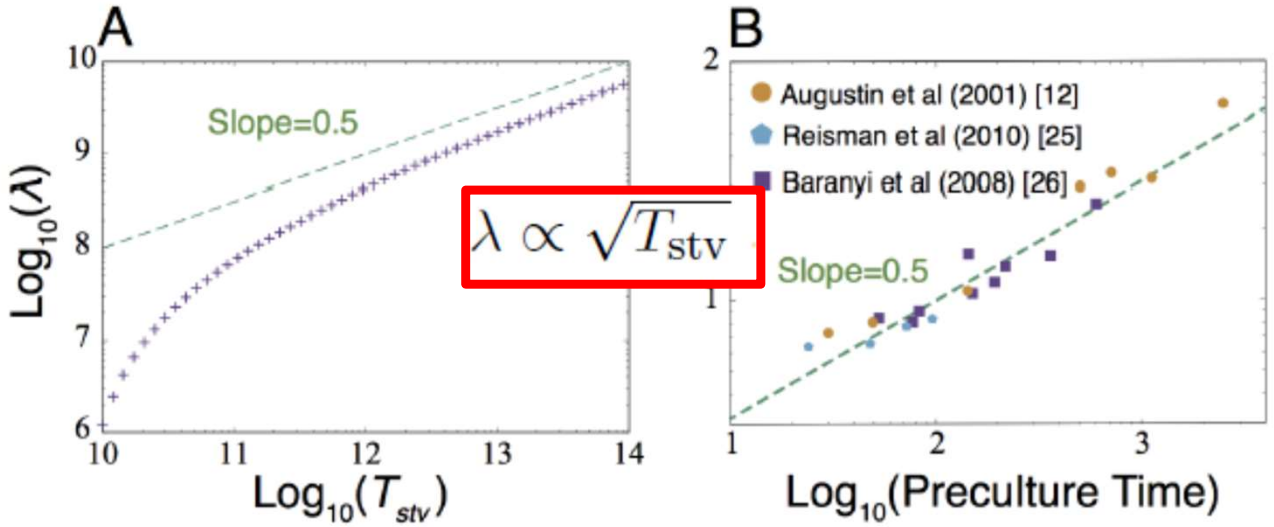
Active Proteins are protected



Transition from exponentially growing state to suppressed growth state (growth rate reduced to 5-6 digits)

→ Waste Inhibits the growth (and degradation) by forming a Complex

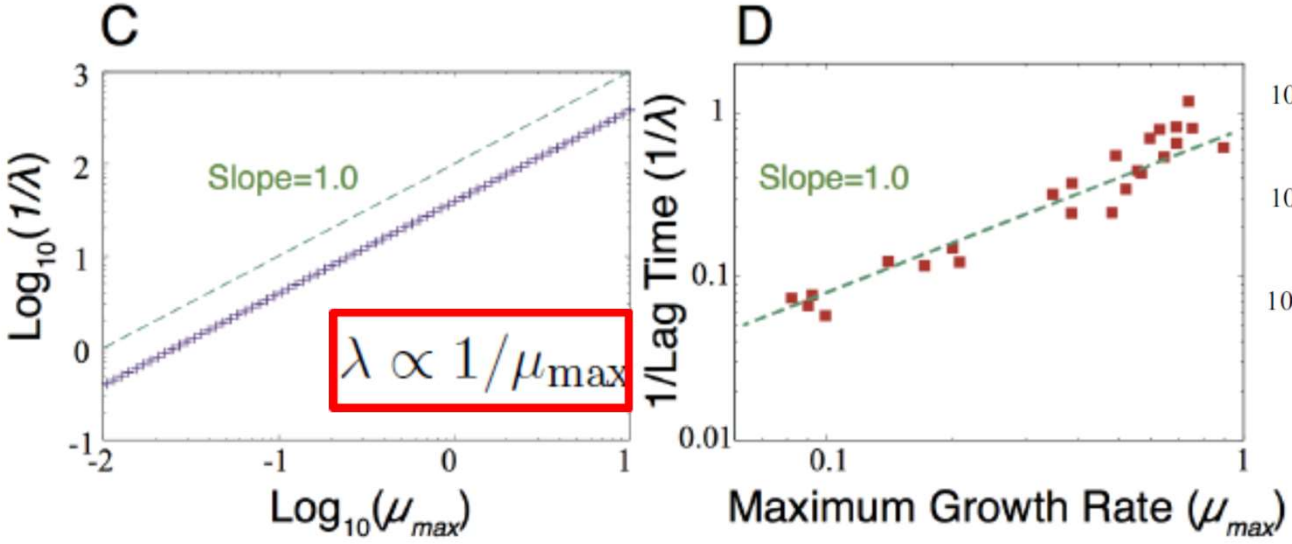
Lag-time λ (needed to recover the growth after resource resumption) follows universal law ～agrees with experiment

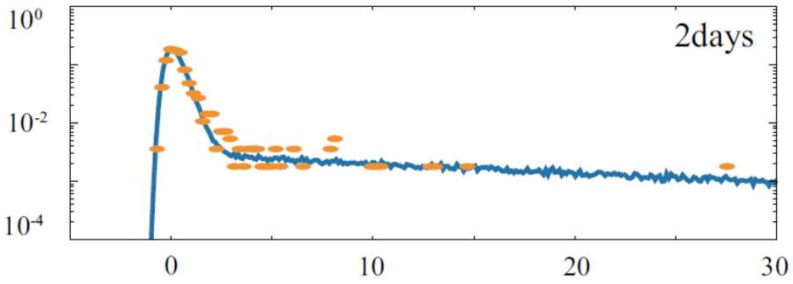


λ ∝ Square-root of starvation time

←dB/dt∝A∝C/B

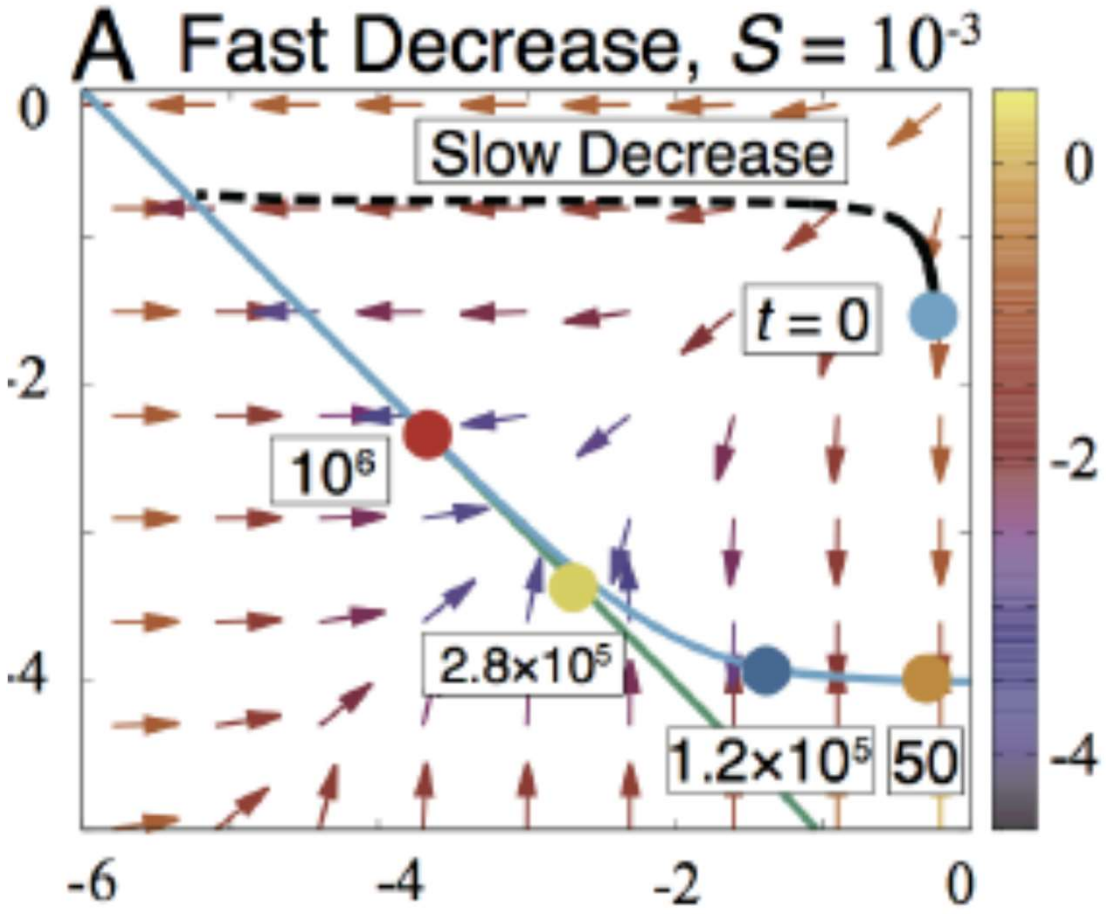
Lag-time distribution by cells
In agreement with experiments





Protocell that can sleep ?

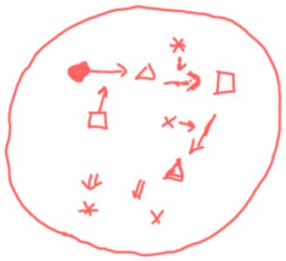
Dynamical-systems mechanism for slow process

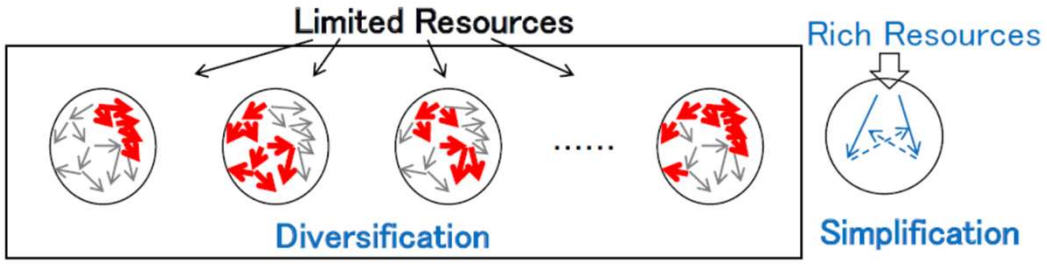


Basic Issues in Origin of Life System molecules→cell

☑ mutually-catalytic diverse components

--- origin/maintenance of diversity? Kamimura,KK 2019





☑ compartmentalization how? Kamimura,KK 2

☑ origin of information(←minority control?)

number symmetry breaking? KK,Yomo 2002

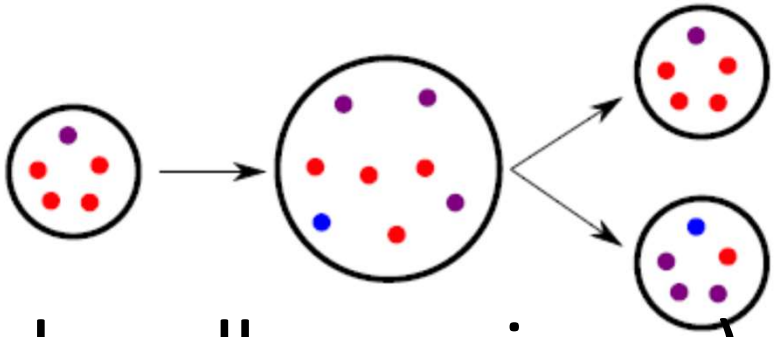
☑ function/information SB:origin of genetic information (central dogma) Takeuchi,KK 2017,2019

Symmetry Breaking from RNA world

Coherent theory? (↑ ???")

- Parasite Question:

Multi-level Evolution



Life is hierarchical: (molecule-cell-organism-..)

Fitness at molecule level ← conflict → cell level

e.g., molecule replication vs cell reproduction

cell replication vs multicellular reproduction

Parasite problem

Evolution of mutually catalytic molecular system

catalytic activity ↑ → replication probability of itself ↓

(← while it forms a complex, it cannot be replicated)

But catalytic activity ↓ cell reproduction rate ↓

Scaling Relation in Multilevel Selection

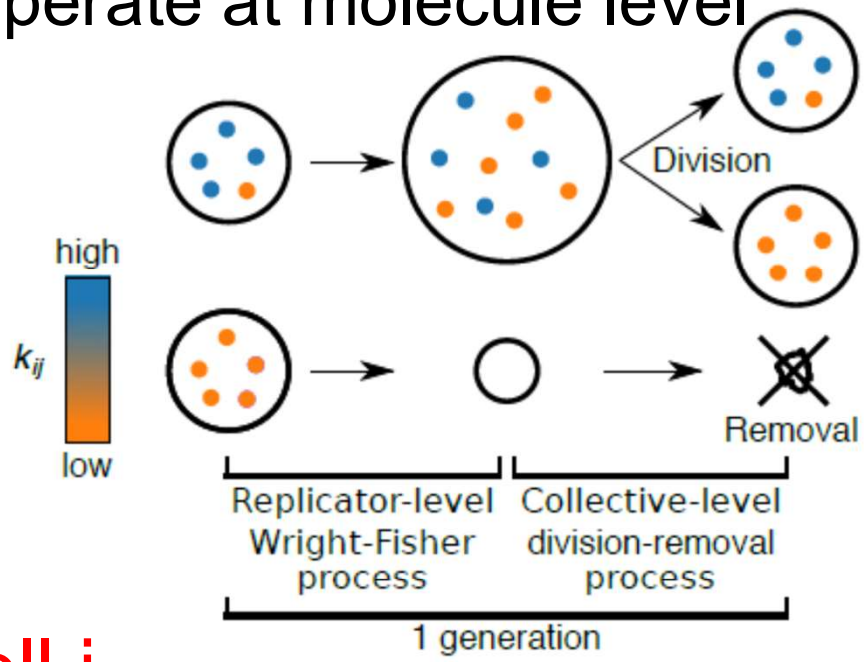
Takeuchi, Mitarai, KK; arXiv 2020

Large N , m(mutation rate) → molecule level dominates (cheaters)
Small N, m → cell level winds (cooperate at molecule level

Boundary (N,m) scaling?

Simplified model

$$w_{ij} = e^{s_a \langle k_{i\bar{j}} \rangle} \frac{e^{-s_w k_{ij}}}{\left\langle e^{-s_w k_{i\bar{j}}} \right\rangle},$$



w_{ij} : Fitness of molecule j at cell i

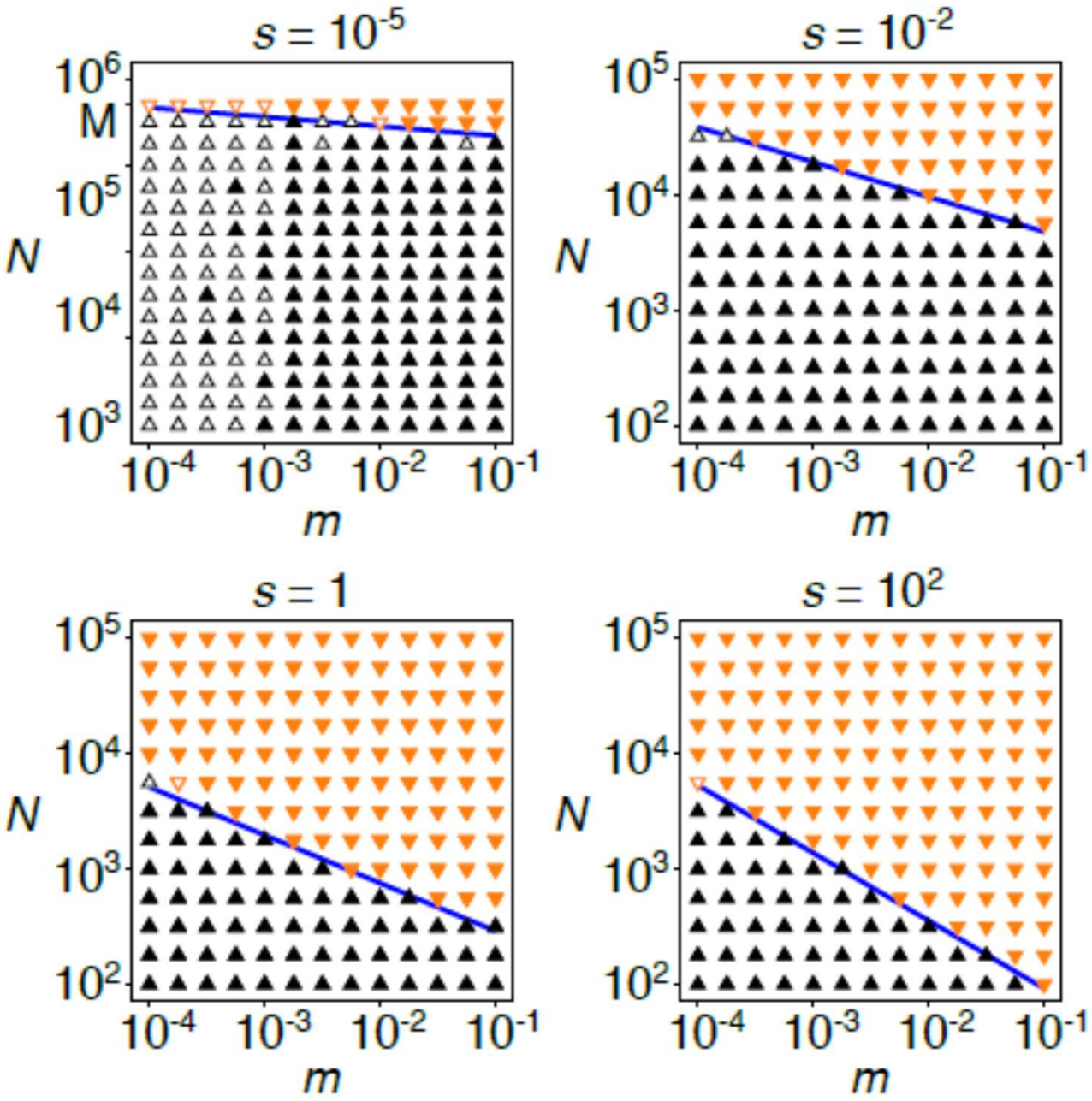
k: molecular activity

FIG. 1. Schematic of model. Replicators (dot) are groupe into collectives (circles). k_{ij} represents amount of public goo replicators provides within collectives.

Cell level (ave. over molecules): larger $\langle k_{ij} \rangle_j$ better
Molecule level smaller k_{ij} better

S: selection pressure

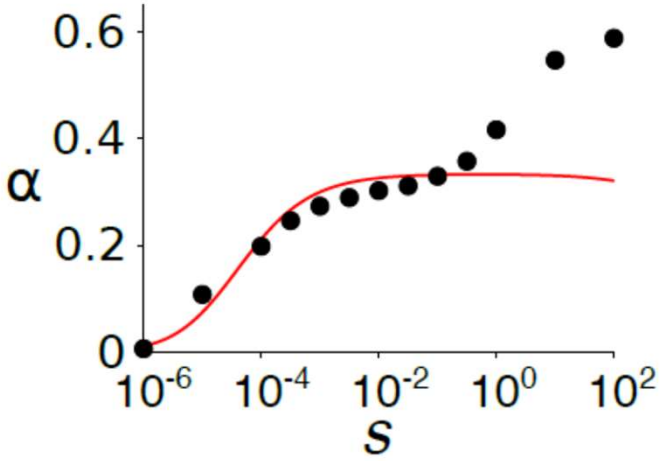
(Fig. 1). Replicator j in collective i is assigned a continuous trait k_{ij} representing the amount of public good



▲ cells grow

▼ activity lost

$N \sim m^{-\alpha}$



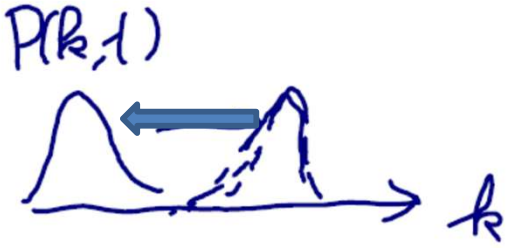
2. Phase diagrams ($M = 5 \times 10^5$ and $\sigma = 10^{-4}$). Sym- have following meaning: $\Delta\langle\langle k_{ij}\rangle\rangle > 3 \times 10^{-7}$ (filled tri- up); $\Delta\langle k_{ij}\rangle < -3 \times 10^{-7}$ (filled triangle down); RHS of

M: total number of molecules in the system

- Theoretical Estimate

molecular level (Fokker-Planck type)

fitness w, activity k



$$\frac{\partial P(k,t)}{\partial t} = m \frac{\partial^2}{\partial k^2} w(k) P(k,t) + (w - \langle w \rangle) P(k,t)$$

$$\langle w \rangle = \int w(k) P(k,t) \, dk$$

<..> average within a cell

$$\Rightarrow \quad \frac{d\langle w \rangle}{dt} = \langle (\delta w)^2 \rangle$$

$$\frac{d\langle (\delta w)^2 \rangle}{dt} = m + O(\langle (\delta w)^3 \rangle$$

Cell level (per generation n; a la Fisher)

$P(\hat{w})$: $\hat{w}$: fitness as a cell

$$P_n(\hat{w}) \longrightarrow P_{n+1}(\hat{w}) = \frac{\hat{w} P_n(\hat{w})}{\int \hat{w} P_n(\hat{w})\, d\hat{w}}$$

$$\Rightarrow \Delta \langle \hat{w} \rangle = \frac{\langle (\delta \hat{w})^2 \rangle}{\langle \hat{w} \rangle} \cdot s$$

$$\Delta \langle (\delta \hat{w})^2 \rangle = s \frac{\langle (\delta \hat{w})^3 \rangle}{\langle \hat{w} \rangle} - O\left(\frac{\langle \delta \hat{w}^2 \rangle^2}{\langle \hat{w} \rangle^2} \right)$$

3rd (central) moment $\sim \gamma \langle (\delta \hat{w})^2 \rangle^{3/2}$

- Now, k (activity) is negatively correlated with molecular replication rate, and <k> is positively correlated with cell reproduction rate

(or one can use (extended) Price eqn)

for the variances of k ($\mathbf{v}_m$ molecular, $\mathbf{v}_c$ cellular)

$$\Delta \mathbf{v}_m = m - \mathbf{v}_m/N$$

Due to sampling in a finite population

$$\Delta \mathbf{v}_c = \mathbf{v}_m/N - \gamma s \mathbf{v}_c^{3/2}$$

Molecular variance is added into cell level

$$\Delta\langle k\rangle \propto (-\mathbf{v}_m + \mathbf{v}_c)s$$

→balance→**V**c=**V**m

Steady-state

$$\mathbf{v}_m = mN$$

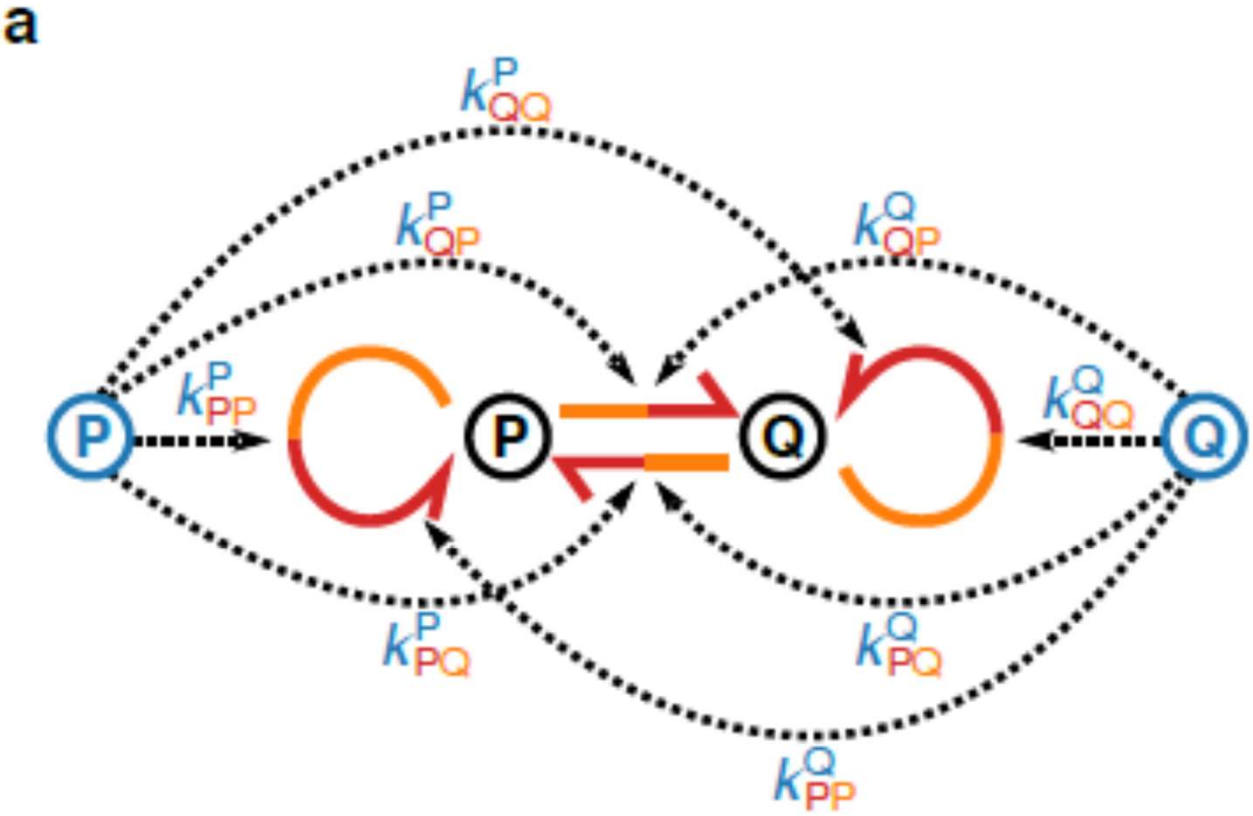
$$\mathbf{v}_c = (\mathbf{v}_m/\gamma sN)^{2/3} = (m/\gamma s)^{2/3}$$

$$\mathbf{v}_m = \mathbf{v}_c \rightarrow \mathbf{m}^{1/3}N = \text{const.}$$

Origin of 'Central Dogma' (2x4 Model)

Two-species (P,Q):initially work as template and catalysts

→Symmetry breaking to Functional (catalytic) vs information (template)molecules through Evolution



meaning	tail	arrow	head
reaction	template	——→	product
catalysis	catalyst	·····→	reaction

k ; catalytic activity for each reaction

In replication – mutation to k+σ(random)

They are in a protocell

V molecules

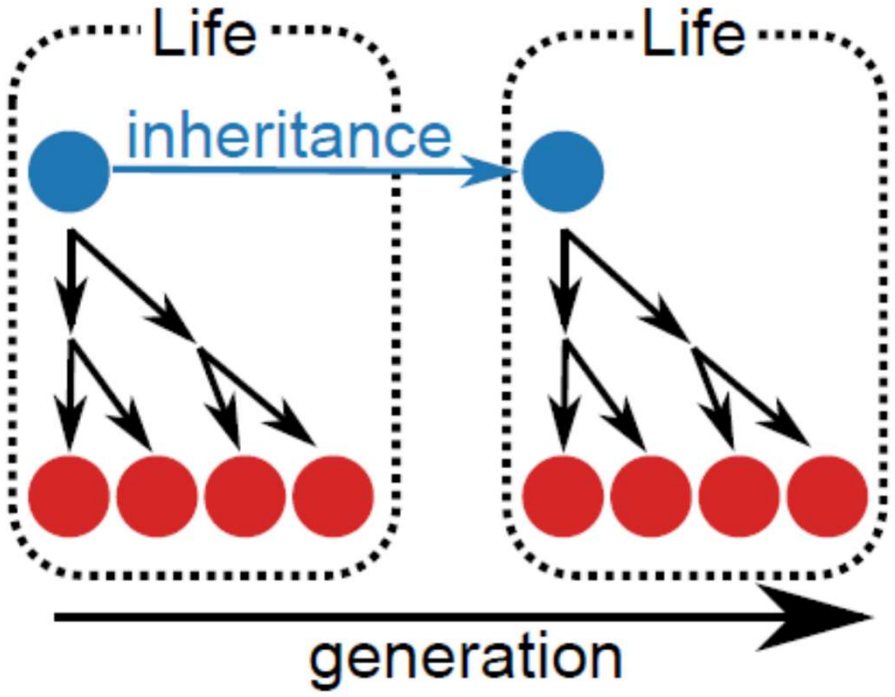
Multilevel selection

Takeuchi, KK, ProcRoySoc 2019

General in conflicting multi-level evolution (?)

Universal feature of life

Hierarchy		Differentiation	
whole	parts	informatic	operational
cell	molecule	genome	enzyme
individual	cell	germ	soma
society	individual	queen	worker



Differentiation for high mutation rate, large

←→ Hamilton's Kin selection theory **? ?**

- 6. Question: Why diversity?
- All life system we know consist of diverse components to be maintained and reproduced

←→ ?Simpler system reproduce faster?

(i) To cope with diverse environmental changes?

(ii) In the beginning complex, just because more probable Complexity in the beginning (Dyson 84)

(iii) To cope against parasitic processes…

(iv) Competition for diverse, limited resources among individuals → 'diversity transition'

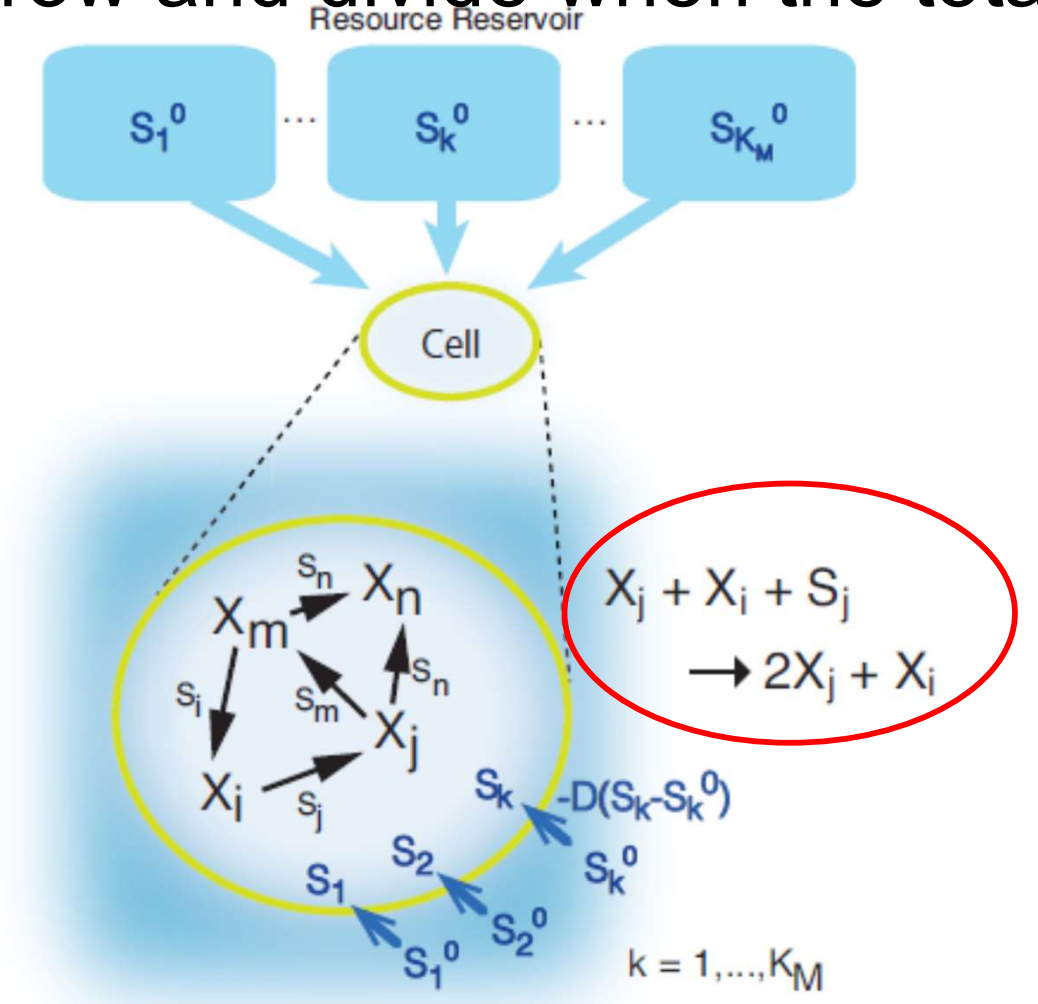
(dXi/dt=AiXi → dXi/dt=Ci)

Complex catalytic network (hypercycle)

Molecular replication vs Cell reproduction

Cell --- chemicals X1,X2,…Xn replicate with the aid of others (hypercycle)

Grow and divide when the total number of molecules=N



Cf: Hypercycle introduced by Manfred Eigen for the issue of the origin of life (i.e., stable replicating system to resolve the error catastrophe)

Simplest Illustration of Diversity Transition

$$X_i + S_i \rightarrow 2X_i (i = 1, ..., K_M)$$

concentration of X_i as $\rho_i (i = 1, ..., K_M)$,

reaction rate a_i
resource S_i

$$\frac{d\rho_i}{dt} = a_i S_i \rho_i - \rho_i \phi, \qquad \phi = \sum_j a_j S_j \rho_j.$$

(Constraint that sum of Xi=1)

$$\frac{dS_i}{dt} = -a_i S_i \rho_i + D_r (S_i^0 - S_i), \quad \Rightarrow \quad \bar{S}_i = \frac{D_r S_i^0}{a_i \rho_i + D_r}. \qquad (1)$$

Rich resource

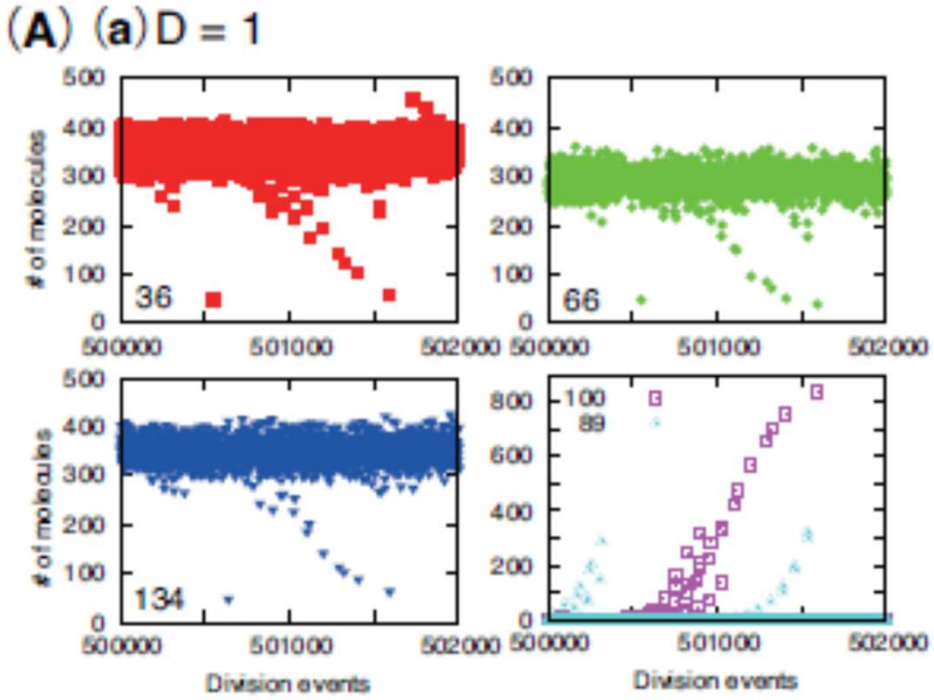
$D_r \gg a_i \rho_i$: ⇒ $\frac{d\rho_i}{dt} = a_i S_i^0 \rho_i - \rho_i \phi_L$, ⇒ Only species with the highest aiSi remains (Darwinian selection)

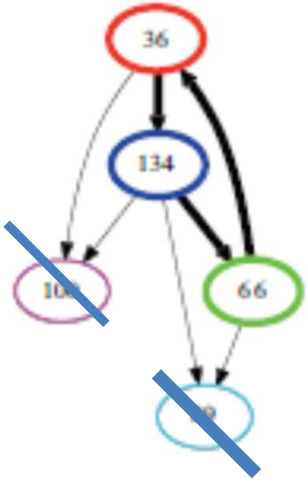
Resource limitation

$D_r \ll a_i \rho_i$ ⇒ $\frac{d\rho_i}{dt} = D_r S_i^0 - \rho_i \phi_S$, ⇒ $\rho_i = S_i^0 / \sum_j S_j^0$,

(Coexistence)

Kamimura,Kaneko,2016





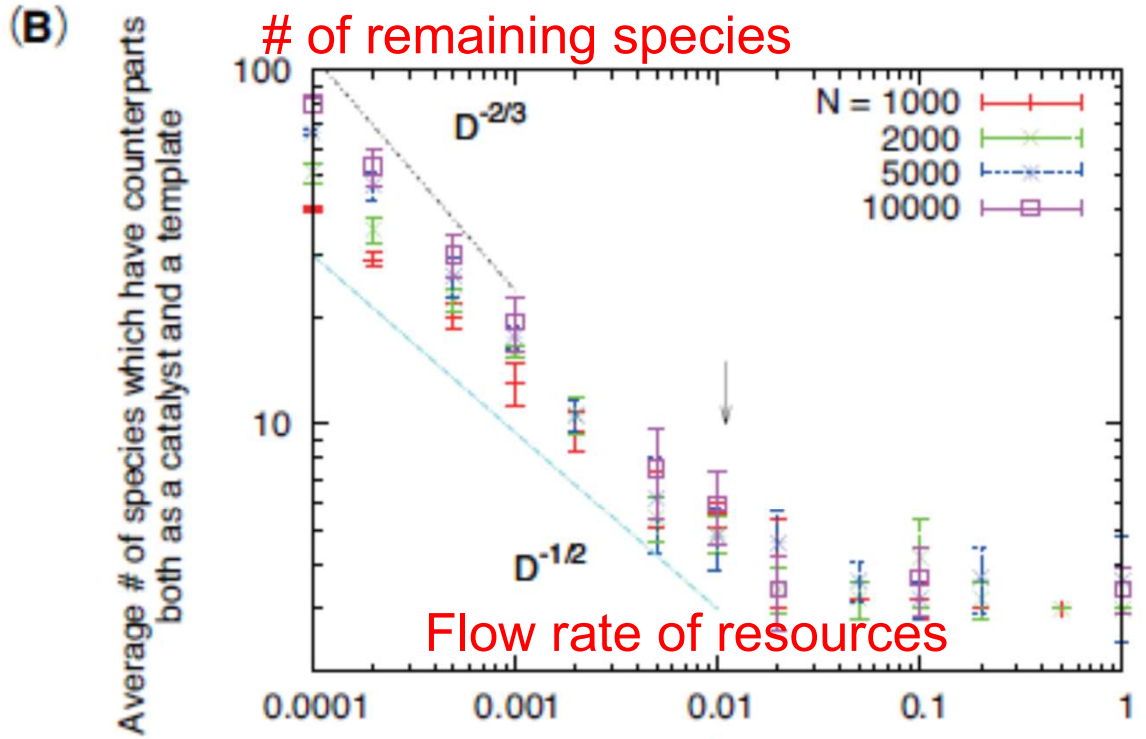
When the resource flow is sufficient, most efficient ,simple hypercycle remains (3 species)

If the resource flow is limited, diverse components remain to form hypercycle-networks

A) Diversity transition when resource flow goes below a threshold ← simple argument

B) Negative Scaling relation between the diversity and resource flow

Reaction rate xi*xj → $(1/K_M)$ $(1/K_M)$ assuming concentration is equally distributed among remaining species K_M



$$\frac{d\rho_i}{dt} \sim \frac{1}{K_M^{*2}} S_i.$$

$$\bar{S}_i = \frac{D_r S_i^0}{1/K_M^{*2} + D_r}. \quad \leftarrow (1)$$

$$G = \sum_i \frac{d\rho_i}{dt} \sim \frac{K_M^* D_r S^0}{1 + D_r K_M^{*2}}.$$

$$dG/dK_M^* = 0 \text{ as } K_M^{\text{opt}} = D_r^{-1/2}.$$

Kamimura,KK,2015,J Systems Chemistry +2016PRE

Q* Propose your own "ideal cell model" that captures some basic property that you think is important, and make some analysis

- (e.g.）metabolic, catalytic process
- Membrane? Information template?
- Postulate/Consequence of consistency among processes?

*** Taking into energetic more seriously?**
resource ーー catabolic（異化）
ーーanabolic（同化）
Amino acid -> ribosome ->protein synthesis..
Enzyme → membrane
DNA? ATP? **（cf Klumpp,Zhang,Hwa(Cell2009)、Youk,Oudenaarden 2009)**
Including waste chemicals?