Complex Systems Biology

cf. Life as Complicated System: (current trend) Enumeration of molecules, processes detailed models describing the life process But understanding??

Life as Complex System:

Understand Universal features at a System with mutual dependence between parts and whole

Strategy:

 1) extension of Dynamical Systems & Statistical Physics
 → Catch consistency between micro-macro levels
 2) Constructive Approach: (Exp & Theory)
 `construct simple system to catch universal features' as simple as possible Consistency between different levels (1)Cell reproduction vs molecule replication (03-) (2)Reproduction of multicellular organism vs of cells (97-(3)Adaptation vs Reproduction (06-) (4)Genetic change vs Phenotypic Fluctuation (03-



Constructive Biology Project

theme	experiment	theory	question
replicating system	in vitro replication with enzymatic reaction	minority control	origin of heredity; evolvability
cell system	replicating cell with internal reactions	universal statistics in reaction dynamics	condition for recursive growth
cell differentiation. development	differentiation of E Coil by interaction	emergence of differentiation rule from dynamics	irreversibility robustness
Spontaneous adaptation	Artificial gene network	Adaptive attractor selection by noise	Robust adaptation without signalling
evolution	Relevance of phenotypic fluctuation and dynamics	Genetic assimilation of phenotype fluct.and dynamics	geno-pheno type relationship

Complex Systems Biology Project (JST, ERATO; KK, Yomo,...)

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



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

C.Furusawa & KK, PRL2003

k species of chemicals , $X_o...X_{k-1}$ number --- n_0 , $n_1 ... n_{k-1}$

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 = n_i, when it exceeds N_{max} the cell divides into two



dX1/dt X0X4; rate equation; Stochastic model here

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

$$dn_i/dt = \sum_{j,\ell} \operatorname{Con}(j, i, \ell) \epsilon n_j n_\ell / N^2$$

-
$$\sum_{j',\ell'} \operatorname{Con}(i, j', \ell') \epsilon n_i n_{\ell'} / N^2$$

+
$$D\sigma_i(\overline{n_i}/V - n_i/N),$$

where Con(*i*, *j*, ℓ) 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



Furusawa &KK,2003,PRL

Zipf's Law is observed at D = Dc



(distribution of x: $(x) x^{-2}$)

Confirmed by gene expression data





Later confirmed by several other groups

Formation of cascade catalytic reaction



With conservation law, The exponent -1 is explained Mean-field type (self-consistent) calc.)



1: minority molecules

- 2 : catalyzed by 1, synthesized by resource
- 3 : catalyzed by 2

• Remarks:

(0) Universality

- (1) Evolution to the critical state (with Zipf law) is confirmed numerically
- (2) Evolution to scale-free network appears later as embedding of power-law abundances into network (Furusawa,KK, PRE 2006)
- (3) Self-organization to critical state, if transport of 'nutrition chemicals' is catalyzed by some chemicals (no need for choice of D) (instead of simple diffusion) (Furusawa,KK,2007)

Evolution of Network to satisfy Zipf's law? Yes Critical D value depends on connectivity in the network; mutation of network + selection \rightarrow approaches Zipf's law



Zipf's law holds, irrespective of network structure, but Later, the connectivity in the network approaches "scale-free" network through evolution. statistical properties; embedded into network structure Dynamics (abundance) first, structure (equation for dynamics) later





Experiment; protein abundances measured by fluorescence



Furusawa,Kashiwagi,,Yomo,KK

Figure 3: The number distribution of the proteins measured by fluorescent intensity. Distributions are obtained from three *Escherichia coli* cell populations containing different reporter plasmids, i.e., EGFP (enhanced green fluorescent protein) under the control of the tetA promoter, DsRed (red fluorescent protein) under the control of the tre promoter with and without IPTG induction. Note that, although the IPTG induction changes the average fluorescent intensity, both the distributions (with and without the induction) can be fitted by log-normal distributions well.

Also studied in GFP synthesis in liposome

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

log(N_x) shows Brownian motion $\rightarrow N_x$ 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 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,



Fluctuations come in parallel:

Usual central limit theorem is valid;

normal distribution.

Replicating artificial cell (experiment) $(\leftarrow \rightarrow \text{theory}; \text{fluctuation}, \text{minority control})$



Tranlation in liposome

Continouos division of liposomes

RNAreplication in liposome

Questions

 (1) All chemicals have such large fluctuations? Important ones are protected?? Origin of heredity (genetic information) why is there genotype and phenotype Minority control mechanism (discreteness important) (KK,Yomo JtheorBiol.2002)

(2)Large phenotypic fluctuation → relevance to biology ?
ans. evolution (Sato et al., PNAS, 2003) adaptation,....
--- recall in standard evolutionary genetics, only the distribution of gene is discussed, by assuming unique phenotype from a given genotype Phenotypic Fluctuation → Relationship to Evolution? selection is based on phenotype (activity, size, protein abundances, fluorescence,...), but

in standard evolutionary genetics; gene a \rightarrow phenotype x uniquely determined

mostly discusses the phenotype distribution as a result of genetic variation

Phenotypic fluctuation of isogenic organisms $\rightarrow P(x; a) x$ —phenotype, a – gene

Artificial selection experiment with bacteria

Selection to increase the fluorescence of protein in bacteria

Schematic drawing of selection process





Organisms with larger phenotypic fluctuation higher evolution speed;

 change of phenotype per generation per mutation --``Response against mutation+selection''

Response $\leftarrow \rightarrow$ Fluctuation

So-called fluctuation-dissipation theorem in physics: Force to change a variable x; response ratio = (shift of x) / force fluctuation of x (without force) response ratio proportional to fluctuation originated by Einstein s paper a century ago...

Generalization::(mathematical formulation) response ratio of some variable x against the change of parameter a versus fluctuation of x

х

P(x;a) x variable, a: control parameter change of the parameter a → peak of P(x;a) (i.e.,<x>average) shifts

$$\frac{\langle x \rangle_{a+\Delta a} - \langle x \rangle_{a}}{\Delta a} \propto \langle (\delta x)^2 \rangle_a = \langle (x - \langle x \rangle)^2 \rangle_a$$

Fluctuation-response relationship (generalized form)

Gaussian distribution of x; under the parameter a

$$P(x; a_0) = N_0 exp(-\frac{(x - X_0)^2}{2\alpha_0}),$$
 at a=a0

Change the parameter from a0 to a

$$\begin{split} P(x:a) &= Nexp(-\frac{(x-X_0)^2}{2\alpha(a)} + v(x,a)) \\ &\quad v(a,x) = C(a-a_0)(x-X_0) + \dots, \text{ with } C \text{ as a constant,} \\ & \longrightarrow \qquad P(x,a_0+\Delta a) = N'exp(-\frac{(x-X_0-C\Delta a\alpha(a_0+\Delta a))^2}{2\alpha(a_0+\Delta a)}) \\ & \text{Hence, we get} \\ &\quad (x > a = a_0 + \Delta a - < x > a = a_0) \\ & \Delta a \\ & \text{Noting that } \alpha = <(\delta x)^2 > \\ & \frac{< x >_{a = a_0 + \Delta a} - < x >_{a = a_0}}{\Delta a} = C < (\delta x)^2 >, \end{split}$$

Artificial selection experiment with bacteria for enzyme with higher catalytic activity for some protein with higher function Change in gene (parameter; a) ``Response'' ----- change of phenotype <x> (e.g., fluorescence intensity) per generation per (synonymous) mutation rate Fluctuation ---- Variance of phenotype x of clone Fluctuation in the phenotype x of clone speed of evolution to increase <x> (proportional or correlated)



Sato, Ito, Yomo, KK, PNAS 2003

• Confirmation by numerical evolution experiment by the reaction-net cell model

Mutate the network ('gene') with mutation rate μ , (rewire the path of the network with the rate) and select such network

having highest concentration c of a specific chemical

- 1. Prepare initial mother cells.
- From each parent cell, mutant cells are generated by randomly replacing reaction paths, with mutation rate µ
- 3. reaction dynamics of all mutants are simulated to determine phenotype x
- 4. Top 5% cells with regard to phenotype x are selected as parent cells of next generation

phenotype $x = log (n_s)$

Confirmation of Fluctuation Dissipation Theorem by reaction-network cell model



(1) the use of log(fluorescence), becauselog x is close to Gaussian distribution in experiments

- (2) New mystery? phenotype fluctuation of clone vs evolution speed in contrast to evolution speed phenotypic fluctuation by genetic variation (Vg): (fundamental theorem of natural selection; established) pheno fluct of clone Vp pheno fluct by gene variation Vg? (fluct by noise variation in 'equation') Follow the spirit of Einstein; gene
- micro-macro consistency \rightarrow Brownian motion



Consider 2-variable distrb P(x=phenotype,a=genotype) =exp(-V(x,a)) Keep a single-peak (stability condition).

KK, Furusawa, 2006 JTB

 $(\partial^2 V/\partial a^2)^{-1} \ge 0; \quad (\partial^2 V/\partial x^2)^{-1} \ge 0.$ $(\partial^2 V/\partial x^2)(\partial^2 V/\partial a^2) - (\partial^2 V/\partial a \partial x)^2 \ge 0.$

Hessian condition

Up to this point pheno (x) and geno (a) are treated in the same way. Then given a, the peak (average) phenotype is x0(a)--function of a --



 $\partial V / \partial x |_{x=x0} = 0$

Gene a is also regarded as variable: Now, we discuss stability P(x,a) as 2 variables instead of P(x;a). write as P(x,a)=exp(-V(x,a))

Condition if it keeps a single-peak solution (stability

annodede concerviont witch, viec cousting, or a board ger on inja imprice

 $\partial^2 V / \partial a^2 \ge 0$ (1)

$$\partial^2 V / \partial x^2 \ge 0$$
 (2)

$$(\partial^2 V/\partial x^2)(\partial^2 V/\partial a^2) - (\partial^2 V/\partial a \partial x)^2 \ge 0$$
(3)

Hessian condition

Up to this point pheno (x) and geno (a) are treated in the same way. Then given a, the peak (average) phenotype is x0(a)--function of a -- $\frac{\partial V}{\partial x}|_{x=x_0} = 0$

x0(a); the average phenotype x of a clone of genotype a; x is given as a function of a Hessian \rightarrow Vg Vip Vg= Variance of x, due to the distribution of gene a

$$\langle (\overline{x}_a - \overline{x}_0)^2 \rangle = V_g = \langle (\delta a)^2 \rangle (\frac{\partial x_0}{\partial a})^2,$$

Vip = Variance of x, for the isogenic organism (i.e., a= a0,fixed)

Vg increases with mutation rate μ , thus, there is critical mutation rate μ c beyond which the distribution becomes flat (error catastrophe).

As Vg μ approximately, Vg = ($\mu / \mu c$) Vip Since Vg evolution speed (Fisher theorem) evolution speed μ Vip \rightarrow fluct-response relation We can do the analysis by using Gaussian 2-body distribution function for phenotype x and gene a; around a=a0, and x=X0;, with coupling between x and a (variance of a is the mutation rate μ

parameter a

For high mutation rate single-peak is not sustained (error catastrophe)



Phenotype fluct. (Vp) vs Gene Fluct. (Vg) in the evolution of toy cell model

Vp: fluct. for given network, Vg: fluct. by network variation



variance of log(x), x is the concentration of the molecule Beyond Darwin with the spirit of Einstein! As µ (mutation rate) increases to µ max,
(1) the distribution collapses (error catastrophe)
(2) evolution no longer progresses beyond µ max evolution speed is maximal at µ ~ µ max
(3) Vg approaches Vp

0.2 mutation rate=0 003 As \mu is increased, mutation rate mutation rate=0.02 The distribution mutation rate=0.03 0.15 mutation rate=0.05 'collapses' requency 0.1 Error catastrophe 0.05 2.82627 32 3.3 2.93.13 nhanahma y

distribution of genotype

- ??? to the theory
- P(x,a) rather than conditional probability (TRICK) "Genetic-Phenotyic correpondence" what phenotype can vary $\leftarrow \rightarrow$ what gene can change fluctuation of variable (micro) vs variation of equation (genetic evolution) (cf Waddington's genetic assimilation) **Q:** Why error catastrophe when Vg>Vip? Robust evolution is possible only under noise -counterintuitive ; it says phenotype noise is important
- → gene-net model

A simple model for Geno-Pheno relationship;
Model:Gene-net(dynamics of stochastic gene expression) → on/off state

Xi - expression of gene i : + (on) - (off) $dx_{i}/dt = \tanh[\beta \sum_{j>k}^{M} J_{ij}x_{j}] - x_{i} + \sigma \eta(t),$ $J_{ij} = -1, 1, 0,$ $(\eta(t)\eta(t') > = \delta(t - t').$ Gaussian white

M;total number of genes, k : output genes

Noise strength

Task

Starting from -1,-1,-1,,-1(all off) xi i=1, 2, ··, k are +1(on) (Target Gene Pattern)

Fitness F= - (Average number of off x_i)

 \dots is temporal average between $t = T_{ini}$ and $t = T_f$

Genetic Algorithm Select networks with higher <F>top--<F>=0Choose top n networks among total N,and mutate with rate μ to produce N networks⁽⁾ (μ :fixed mutation rate)

Result of evolution

Top:reaches the fittest

faster for lower noise(

Lowest; cannot evolve for low noise()





Fitness Distribution

< c --low fitness mutants distributed
 > c - eliminated
 through evolution



Existence of critical noise level c below which low-fitness mutants accumulate (error catastrophe)

(1) Vip Vg for c
(2) Vg Vip as c
(3) evolution progresse only for Vip Vg

(4) Vip Vg through evolution course



Theory confirmed



Why?; difference in basin structure

- > c → large basin for target attractor
 (robust, (distance to basin boudary)
- < c → only tiny basin around target orbit
 remains small



why threshold?

choose paths to avoid turning pts within (noise)

Mutation touches turning points within range of μ

small - > an orbit with small can reach the target



Deviation of basin boundary (turning points) by Noise - > p by Mutation -> g

increases --> robustness increases if g > p, mutation destroys ~d the history ->Vip>Vg necessary for evolution of robustness



~distance to turning points (basin boundary)

Discussion

- Developmental Robustness to noise ---- Vip
- Robustness to mutation in evolution ----Vg
 When Vip>Vg, both decrease, i.e., robustness ✓
 Noise is necessary for evolution of robustness
 Vip Vg → Developmental robustness and genetic

(evolutionary) robustness are linked

Generality of our result; For a system satisfying:
 (1) fitness is determined after developmental dynamics
 (2) developmental dynamics is complex
 (eg.,with distributed catastrophic pt leading to error
 (3) effective equivalence between mutations and noise with regards to the consequence to fitness
 (→ genetic assimilation by Waddington)

Spontaneous Adaptation

- For all possible changes in environment, signal transduction network is already provided?
- Or, is there any general (primitive) mechanism to make spontaneous adaptation?
- → Constructive Experiment with artificial Gene and theory assuming only growth condition and stochsticity

(ex) Adaptive response without signal transduction

Embedded gene network

Unexpected; beyond designed

Selection of preferable state

Phenomenological theory of attractor selection



- Embedded network: each of the two can be selected equally. However, 'good' attractor in each environment is selected. Why?
- Due to hidden signal network?
 NO!: verified by exchanging the promoter
- After each state is attracted with 50%, cells in a 'bad' attractor cannot grow, cells in a good attractor can grow, so that good attractors are selected?
 - NO!; the process occurs without (or before) the cell division process
- Novel Mechanism of Spontaneous Adaptation (without the use of signal transduction) should exist!

• Basic Logic

 $dx_i/dt = F(Activity)f(x_i)-G(Activity)x_i+$ (t)

F,G: increase with activity.

- active: synthesis, degradation both are fast \rightarrow noise
- Active state : both Ff and Gx are large
 - deterministic part >> noise
- **Poor state** : both Ff and Gx are small

deterministic part ~ noise

Switch from Poor state to Active state by noise

(Kashiwagi, Urabe, kk, Yomo; submitted)

The mechanism for adaptive response by attractor selection

$$\frac{d}{dt}m1 = \frac{syn(act)}{1+m2^2} - \deg(act) \times m1 + \eta_1$$
$$\frac{d}{dt}m2 = \frac{syn(act)}{1+m1^2} - \deg(act) \times m2 + \eta_2$$

$$syn(act) = \frac{6act}{2+act}; deg(act) = act;$$

$$\frac{d}{dt}act = \frac{pro}{\left(\left(\frac{Nut_thread_1}{m1+Nutrient1}\right)^{n_1}+1\right)\times\left(\left(\frac{Nut_thread_2}{m2+Nutrient2}\right)^{n_2}+1\right)} - cons \times act$$



Adaptive Response of the genetic network to a environmental change





- Growth-Induced-Attractor-Selection (Furusawa kk)
- Basic Logic $dx_i/dt = f(x_i) - S(\{x_i\})x_i + (t)$ $S \rightarrow$ dilution effect \rightarrow noise Active state : both f and S are large deterministic part >> noise Poor state : both f and S are small deterministic part ~ noise Switch from Poor state to Active state by noise Selection before reproduction General logic in a system with growth and fluctuation



Gene network -> a huge number of attractors coexist with different growth speeds



Spontaneous selection of optimal growth states General in a system with noise and growth Summary

Consistency Principle for Biology

- -- replication of molecules and cells :Universal Laws
- (-- replication of cells and cell ensembles)
- --- genetic and phenotypic changes

(+speciation)

--- adaptation of internal cellular state and growth

·Biological relevance of phenotype fluctuations?

→Phenotypic Fluctuation Evolution Speed

→ Relation between

(isogenic)phenotype fluctuation vs phenotype variation by mutation

- Robustness to mutation and to developmental noise are linked
- Growth system \rightarrow general adaptation by noise
- consequence of steady growth system

Kunihiko Kaneko

D Springer

COMPLEX SYSTEMS

UNDERSTANDING Springer:

Life: **An Introduction** to Complex Systems Biology

Collaborators Chikara Furusawa Katsuhiko Sato

experiment

Tetsuya Yomo **Yochiro Ito** Akiko Kashiwagi

Most papers (biology, **Dynamical systems**) Available at http://chaos.c.u-tokyo.ac.jp

ERATO Complex Systems Biology Project

(2006, August)