Harnessing Complexity
through
Evolutionary Dimensional Reduction

Kunihiko Kaneko

Universal Biology Institute, U. Tokyo
→
Niels Bohr Institute

(My) standpoint in Universal Biology

Life System consists of diverse components,
maintains itself and can continue to produce itself

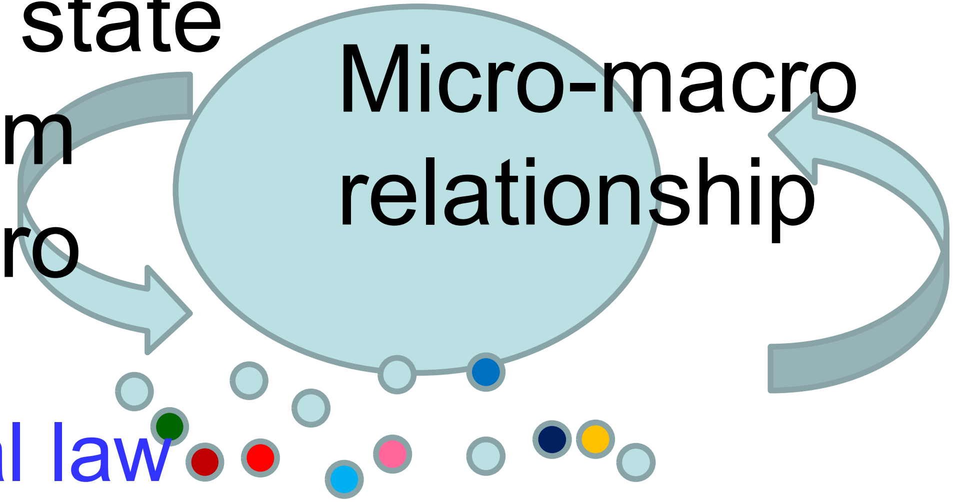
Guiding Principle-- **Macro-Micro Consistency**:

micro – diverse components (high-dimensional)
Thousands of chemical species

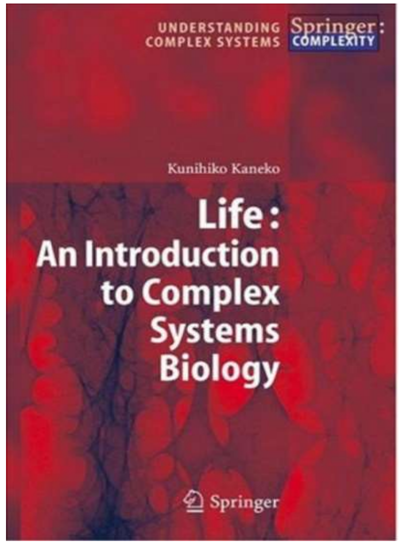
macro – unit to sustain/ reproduce as a whole
(*low-dimensional description*?)

molecule – cell, cell-tissue etc.

Steady (growth) state
Constraint from
macro to micro

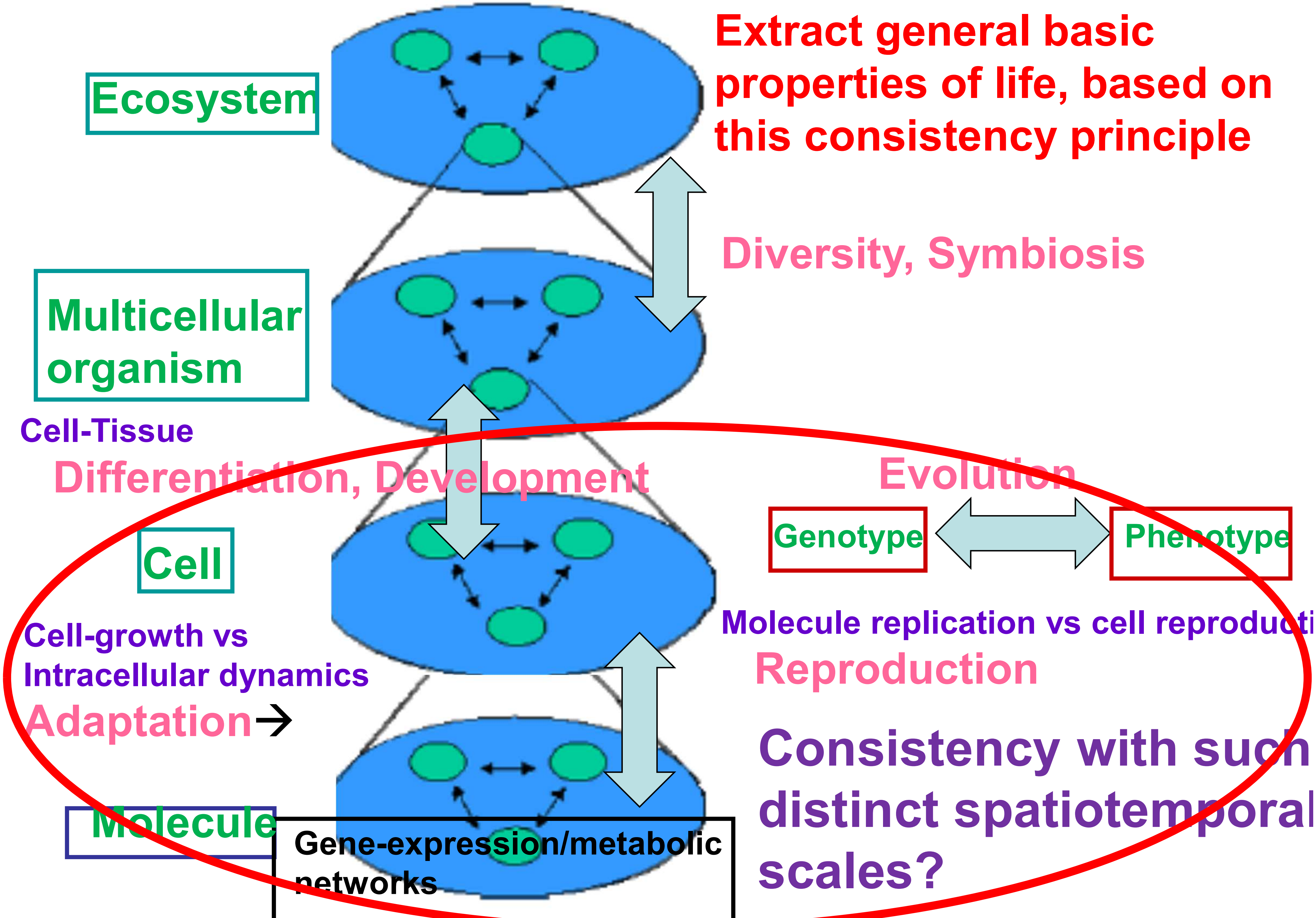


Universal statistical law



Complex-systems
Biology

Consistency between hierarchical levels (+collapse)



Consistency between dynamics of different levels

(1)Cell reproduction vs molecule replication →
universal statistical laws in gene expression
(Furusawa et al, PRL 2003,2012, Biophysics 2006,KK etal, PRX2015)

(2)Adaptation → universal adaptation laws (Kashiwagi et al Plos One2005; Furusawa, KK Phys RevE2018)

(3) Differentiation: **Cell vs multicellularity →**

Oscillatory dynamics => pluripotency + cell-cell interaction → differentiation, loss of pluripotency

(KK&Yomo 1997, Furusawa&KK,1998,Science 2012)

(4) Genetic vs phenotypic changes→

Isogneic Phenotypic Variance by noise ∝ variance by genetic change Vg ∝ Evolution Speed (plasticity)

Robustness to noise ～ to robustness to genetic change, (PNAS03,PLosOne07,Furusawa,KK,Interface2015,PRE 2018)

Part I: Consistency (with robustness) between molecule and cell levels :

→ Evolutionary Dimensional Reduction in phenotypic dynamics

→ Law in Adaptation and Evolution

Response Theory

Part II: Evolutionary Fluctuation-Response Relationship

→Pheno Variance by noise ∝ that by mutation

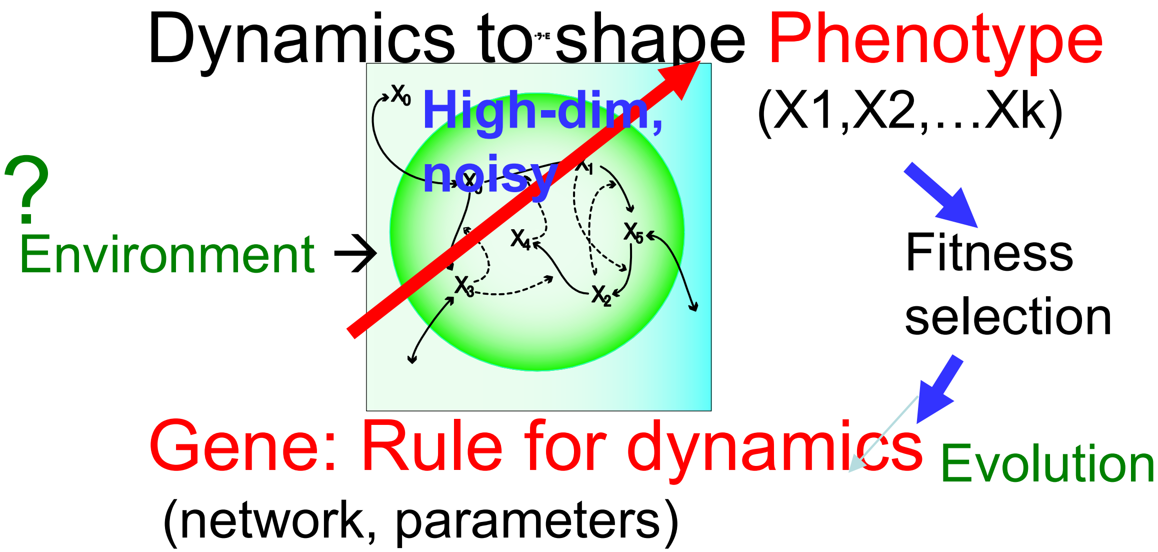
∝ evolution speed

Phenotypic Evolution is directed (predictable), before genetic evolution

- Basic Setup (Exp/Theory/Model)
- Phenotype=Abundances of each component (e.g., protein/mRNA) (~5000 dimensions)

Genotype- DNA seq, or rule for dynamics:

Geno-Pheno Mapping?



* **Experiment:** transcription analysis of E Coli

* **Model:** (i)catalytic reaction network for growth

(ii) Gene regulation net: (high-dim dynamics):

* **Theory**: Low-dim constraint in high-dim states

Trivial(?) Law in Adaptation: Focus on steady-growth cells → universal constraint

all the components have to be roughly doubled (for cell division) : steady-growth condition

Xi – log(concentration of component i) (i=1,,,M)

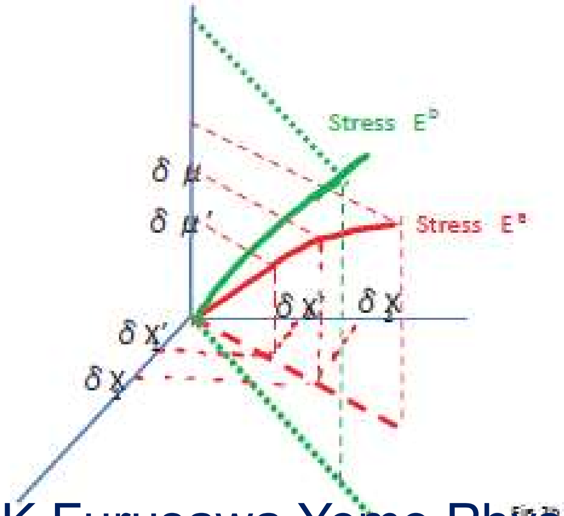
→(M-1) conditions → 1-dimensional line

M large: e.g., # of protein species ~(10^3 ~ 10^4)

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

E: Environment; δE; added Stress

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



Linearization, "small" δE, δX、δμ

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

for given type of stress E (changing strength)

KK,Furusawa,Yomo,PhysRevX(2015)

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

Temporal change of concentration x (Any reaction dynamics)

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

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 $X = log x$, and $f_i = x_i F_i$. Then,

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

Response under different stress strength E

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

Trivial so far

Linearization w.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}. for F({X})

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

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

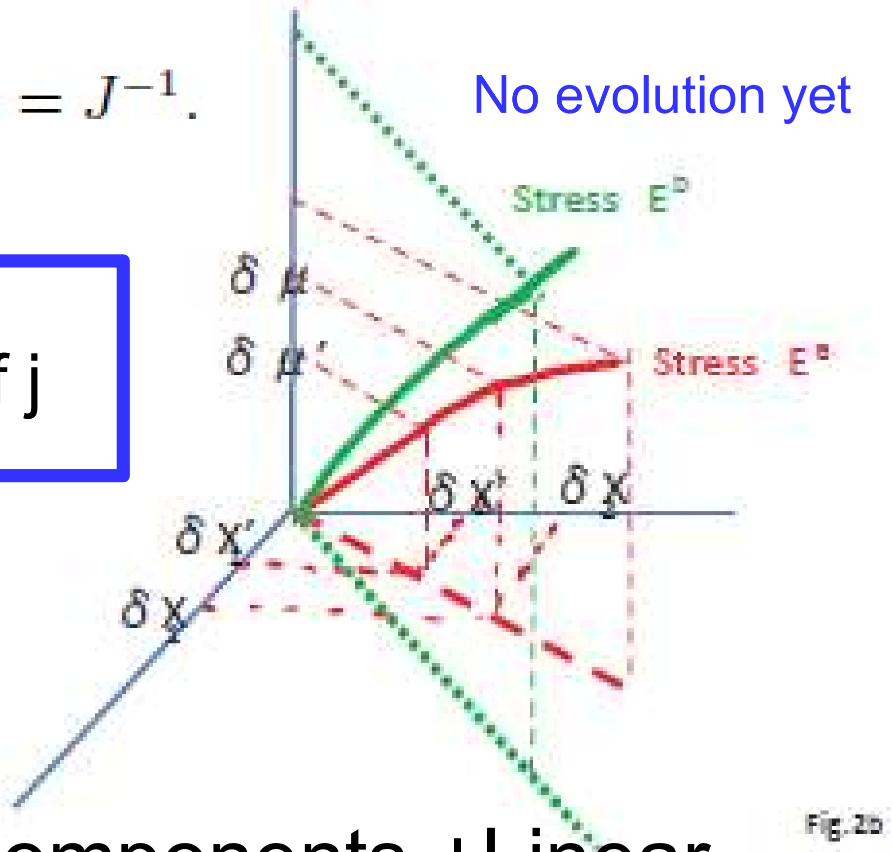
Trivial
+ linearization

No evolution yet

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

$L = J^{-1}$.

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



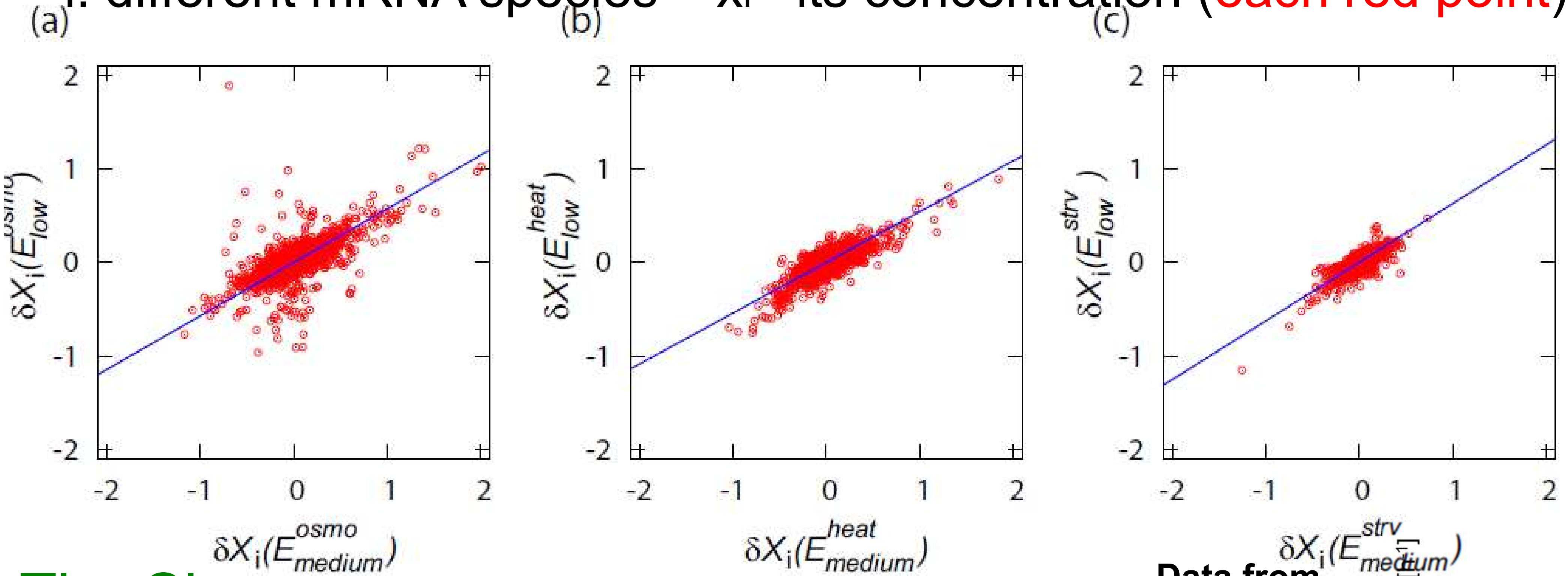
Common proportionality for log-expression change δXj for all components j

←Steady-growth sustaining all components +Linear

Put E Coli under different strength of stresses; Measure gene expressions (mRNA concentrations)

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

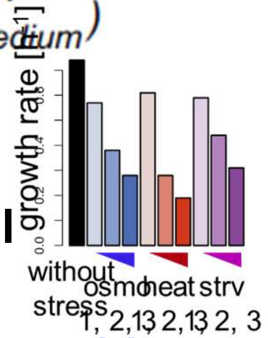
i: different mRNA species x_i –its concentration (each red point)



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 l2013



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

Linearization works for too(?) broad regime

Non-trivial point: Emergent "Deep Linearity"

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

--beyond just steady-growth system

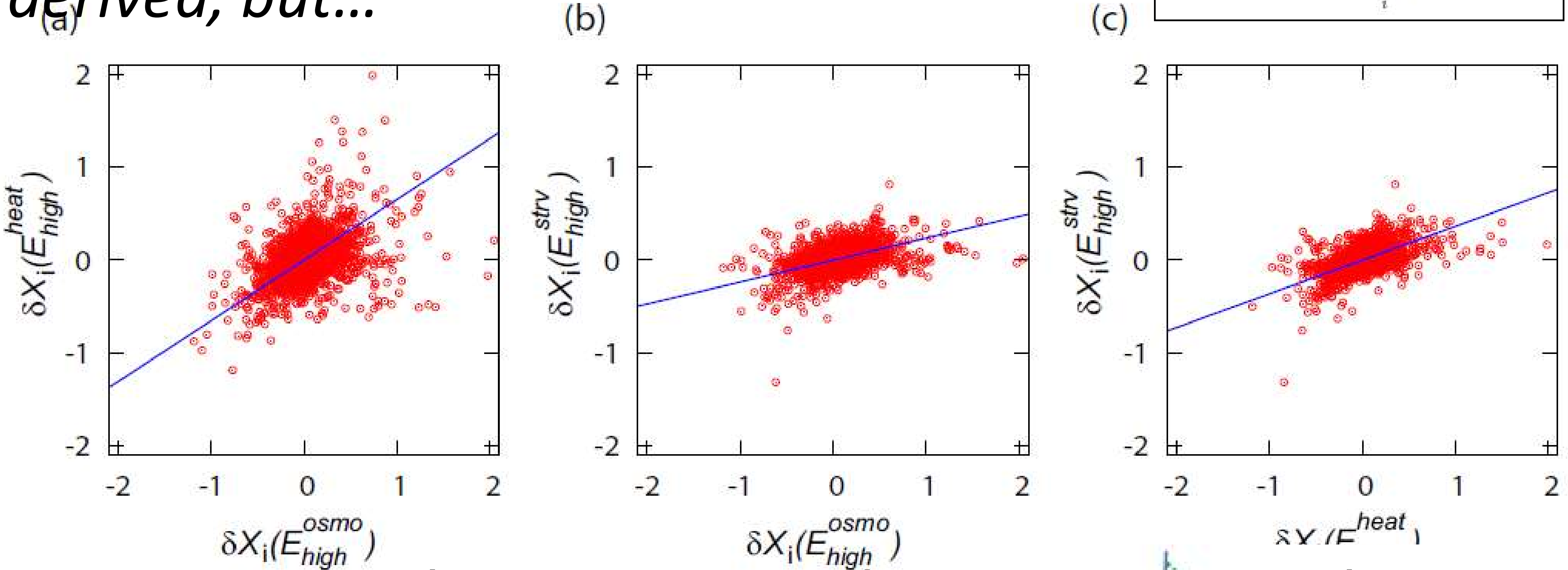
achieved in an evolved system ?

Across Different types of stresses:

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

γi(*a*) depends on stress type *a* *so correlation not derived, but…*

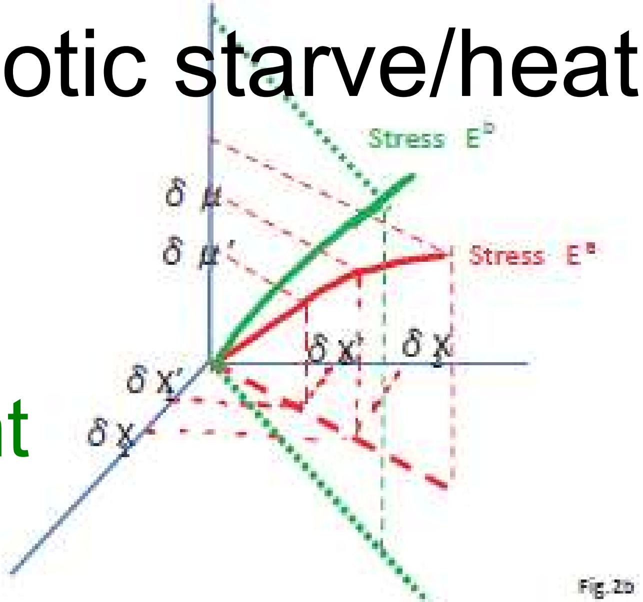
$$\delta X_j(E) = \delta\mu(E) \times \sum_i L_{ji}(1 - \gamma_i/\alpha)$$



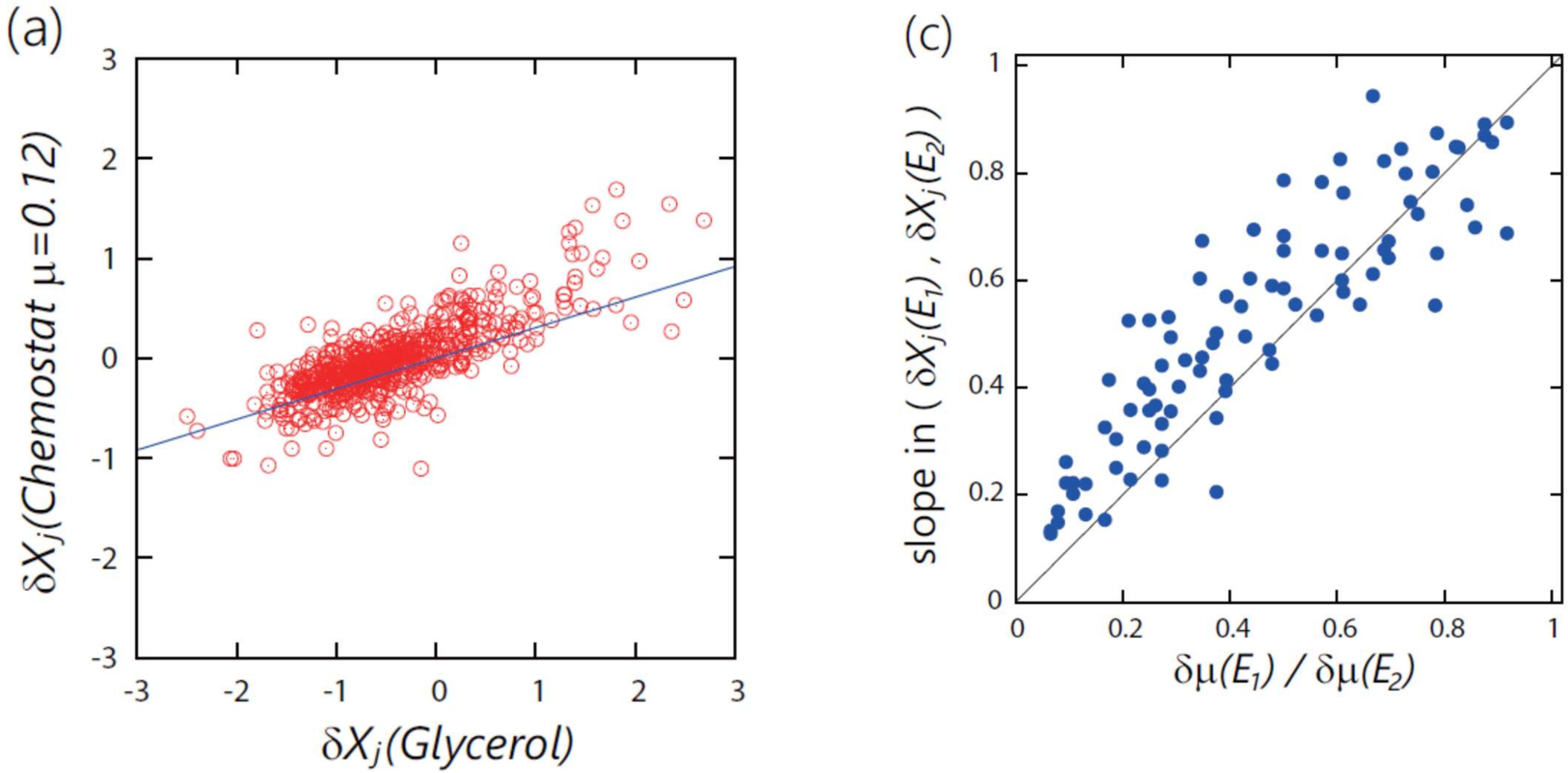
osmotic / heat starve/osmotic starve/heat

Still highly correlated

Confirmed also in protein expression changes across different environmental conditions



Better(?) confirmed in protein expression changes across different environmental conditions (based on the data by Heinemann)
20 different conditions on E Coli

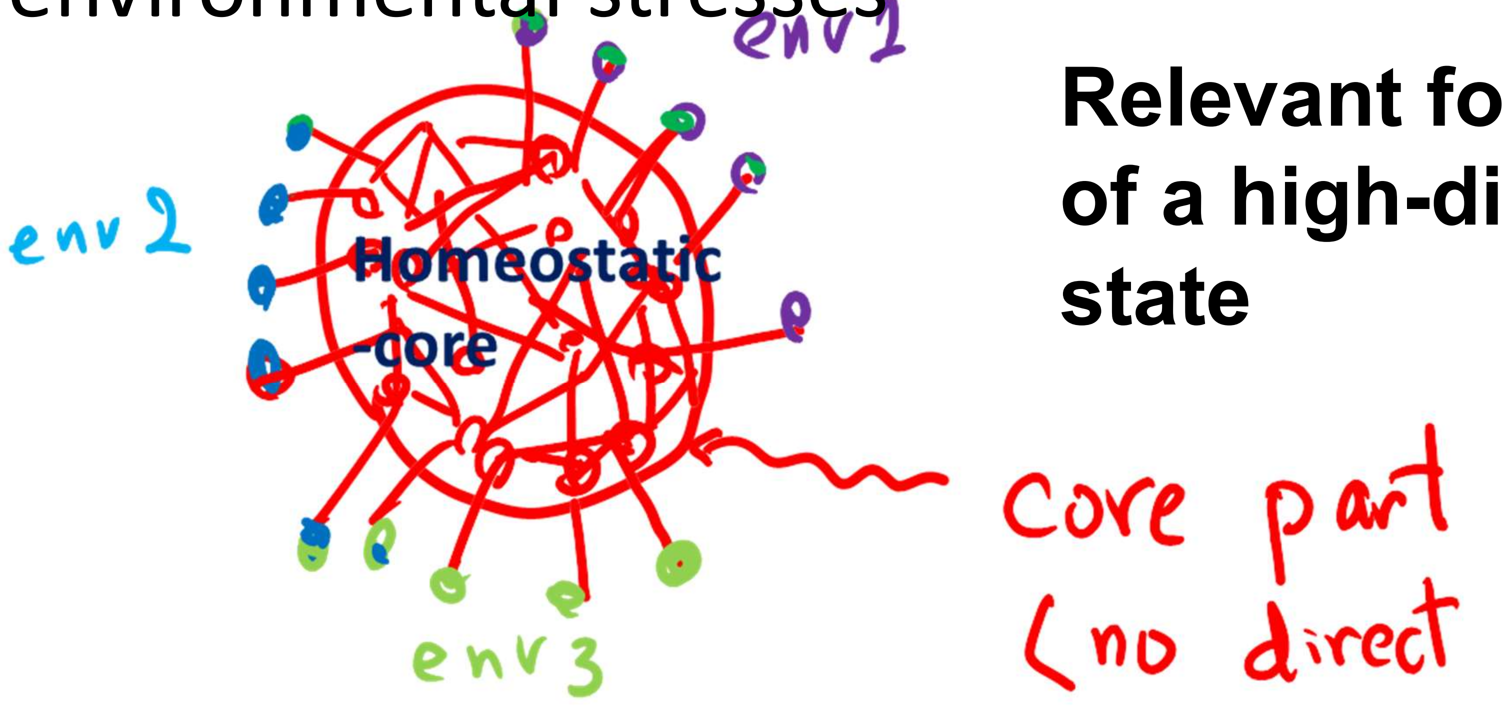


Furusawa, KK Phys.Rev.E 2018

- High-dimensional adaptation system (diversity) is important for expanded liner regime and applicability for diverse environmental changes

  * emergence of ‘collective’ slow variable (Image)

homeostatic core (major parts) -- proportional change, self-consistent ; few genes absorb specific environmental stresses



Relevant for robustness of a high-dimensional state

Non-trivial point: Emergent "Deep Linearity"

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

--beyond just steady-growth system

achieved in an evolved system ?

Check by simulations of toy models with high-dim dynamical systems

Examine by Toy Cell Model with Catalytic Reaction Network

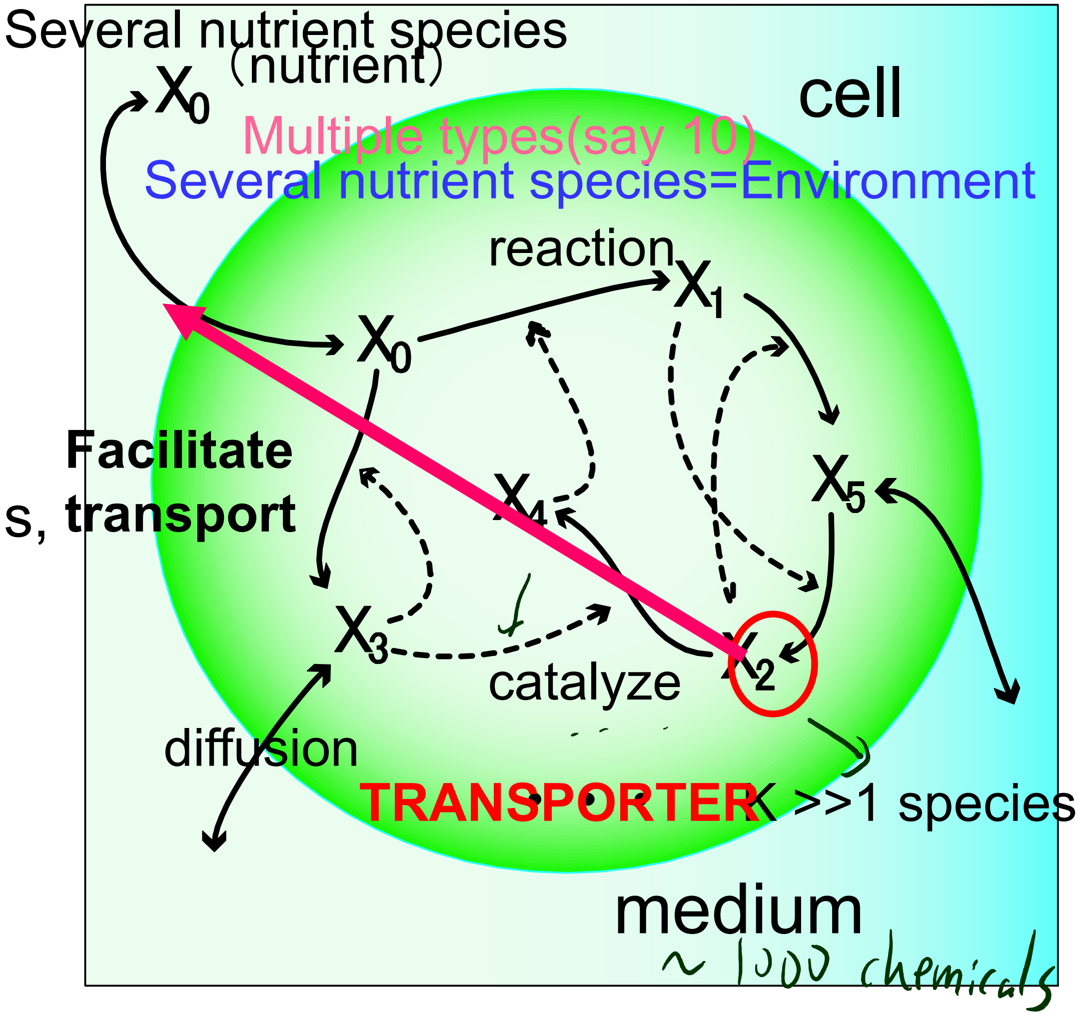
(Cf. Furusawa,KK, PRL 2003, 2012)

- k **species of chemicals** 、$X_o \cdots X_{k-1}$
number ---n_0 、n_1 … n_{k-1}
- random catalytic reaction network
with the path rate p
for the reaction $X_i + X_j \to X_k + X_j$

☐ **Resource chemicals (<- environment) are transported with the aid of a given catalyst, transporter**

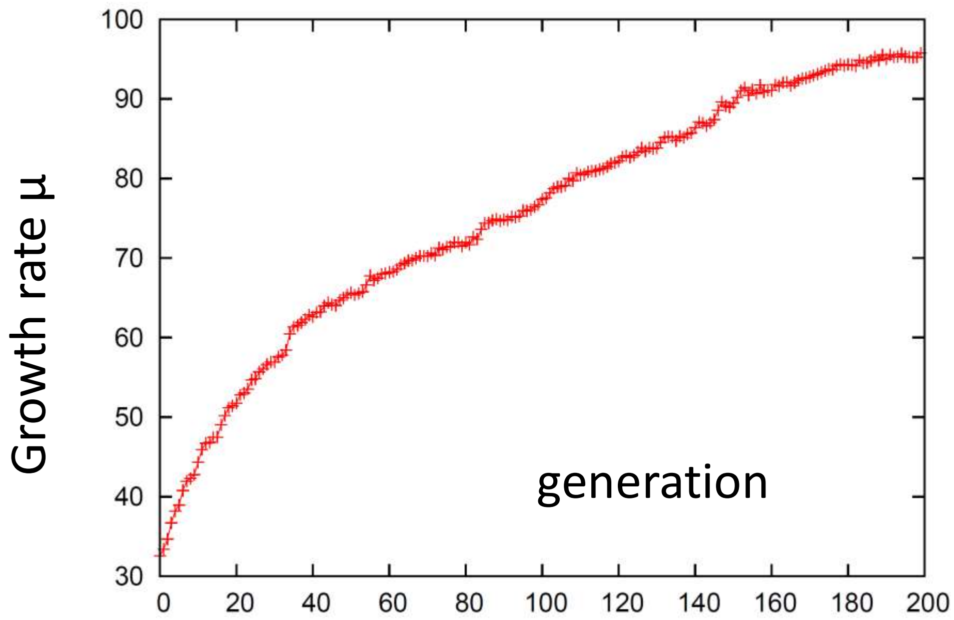
- resource chemicals are thus transformed into impenetrable chemicals, leading to the growth.
- $N = \Sigma n_i$ exceeds N_{max} (model 1)
- Genotype: Network;
- Fitness: e.g., growth rate
- Evolution: Mutate reaction paths, and select those with higher fitness

Model (stochastic reactions)



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

Evolve Network to increase the growth rate under given resource condition



resource concentrations i=1,2,..,10 e.g., (e0,e0,,,e0)

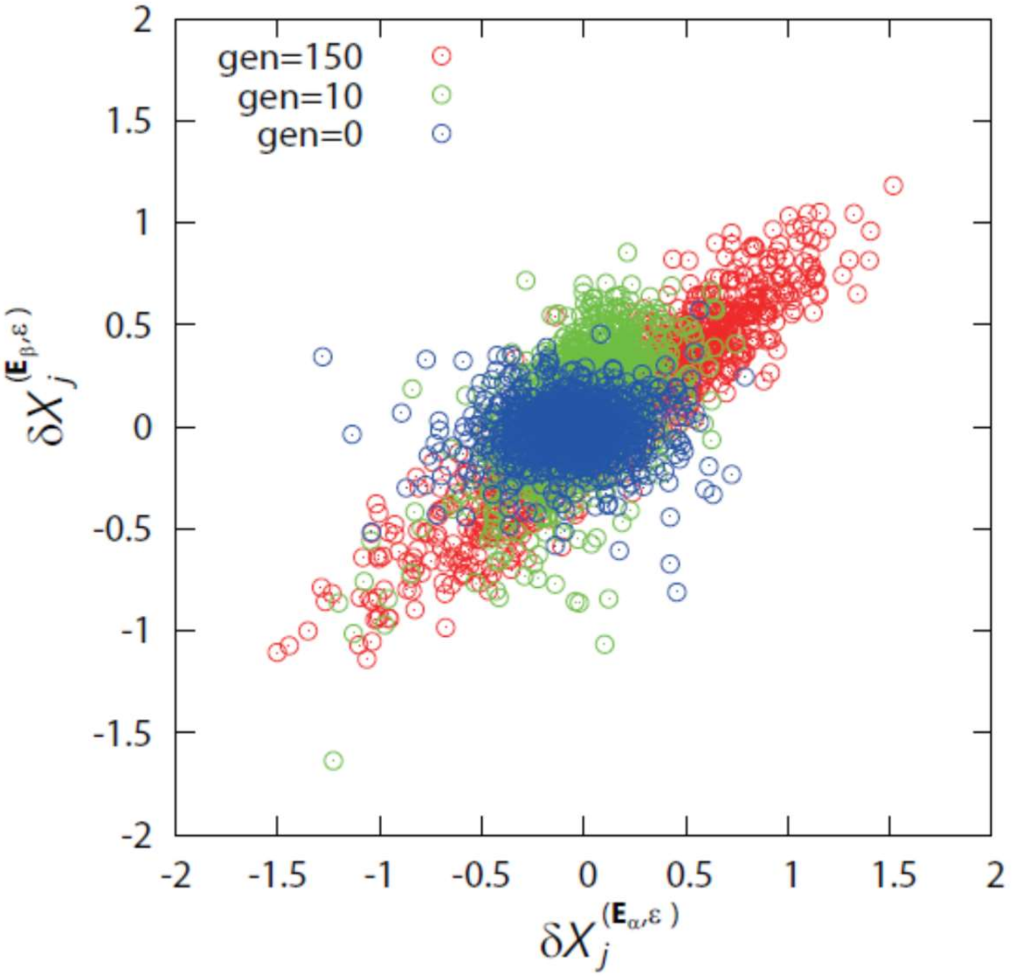
Then put different environmental conditions

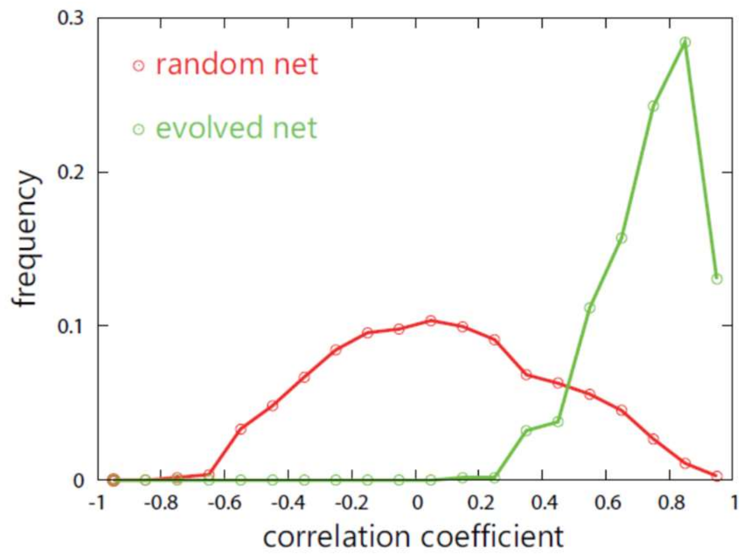
Env = λ （e1,e2,e3,..e10) + (1-λ）(e0,e0,…, e0)

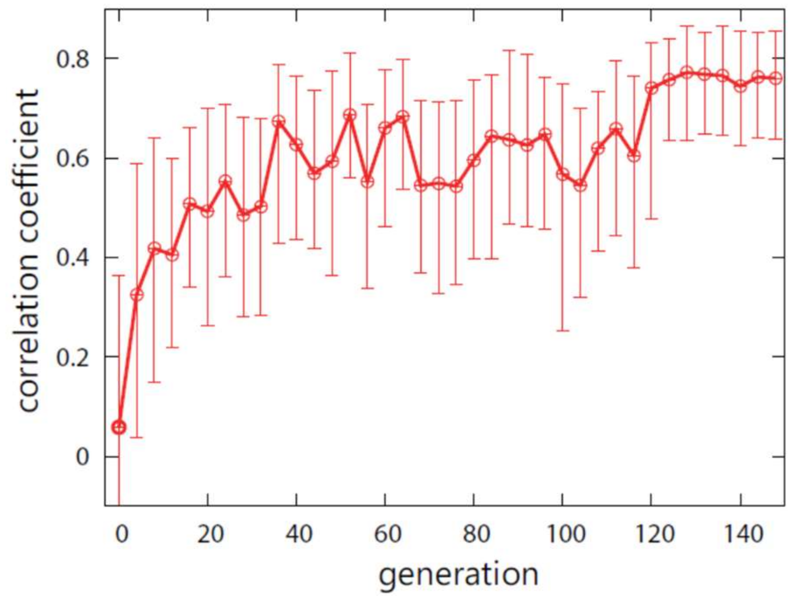
-1< e1,e2,… <1 (randomly chosen)

Check the change in concentrations and growth rates against λ

Evolution shapes Global Proportionality across different environmental conditions







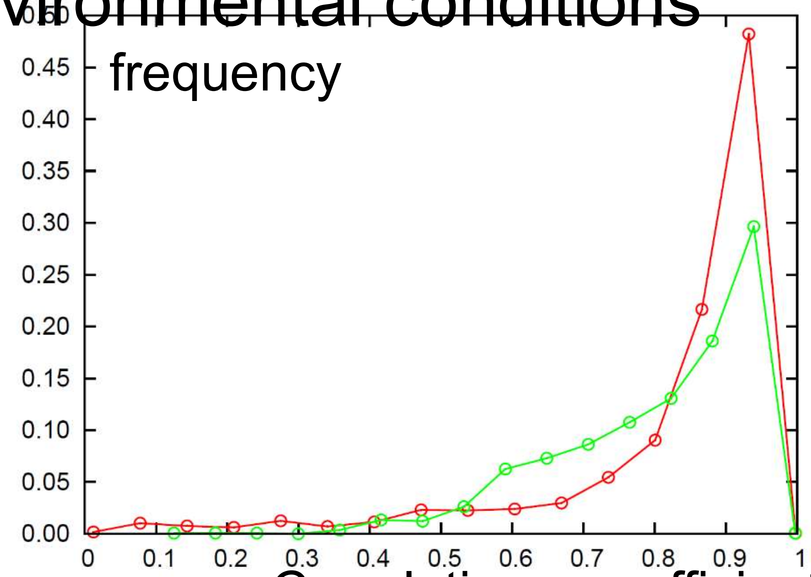
Furusawa, KK, Phys.Rev E 2018
KK, Furusawa, Ann Rev Biophys 2018

After evolution, correlation across different env cond. ☑
Increases + slope-growth-rate proportional

Between same environmental conditions

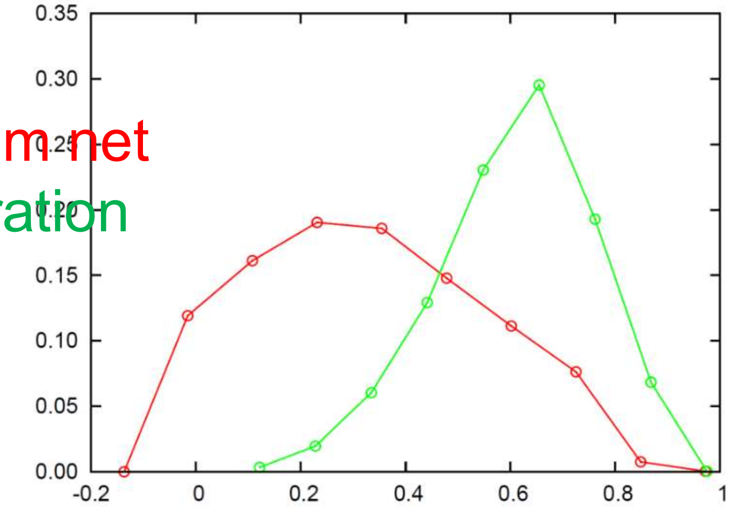
Across different env conditions

frequency



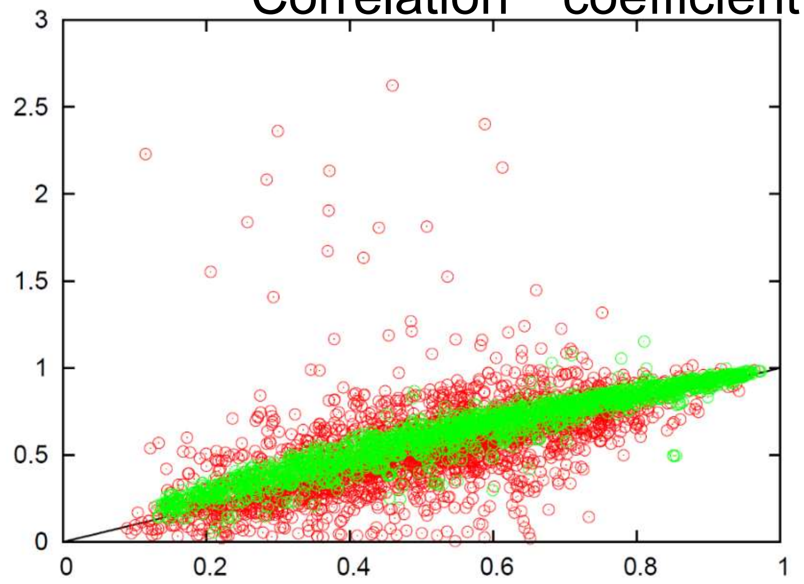
random net

generation =150



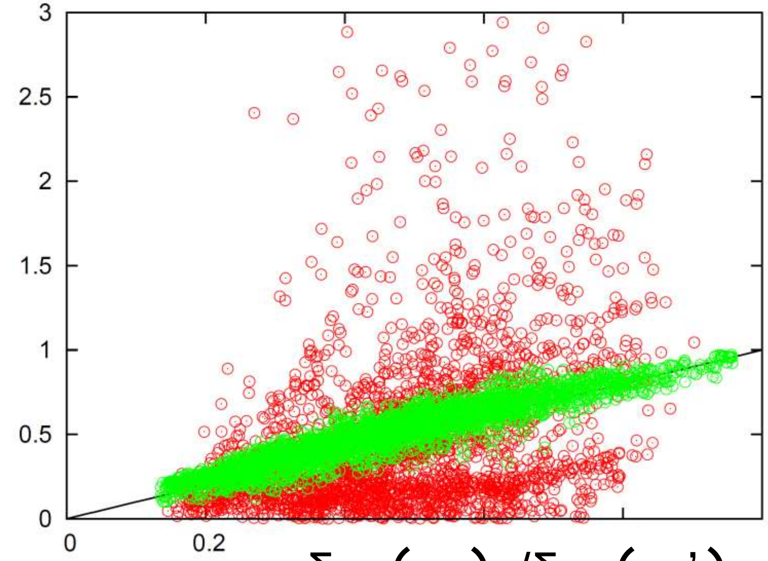
Correlation coefficient across component concentrations

slope



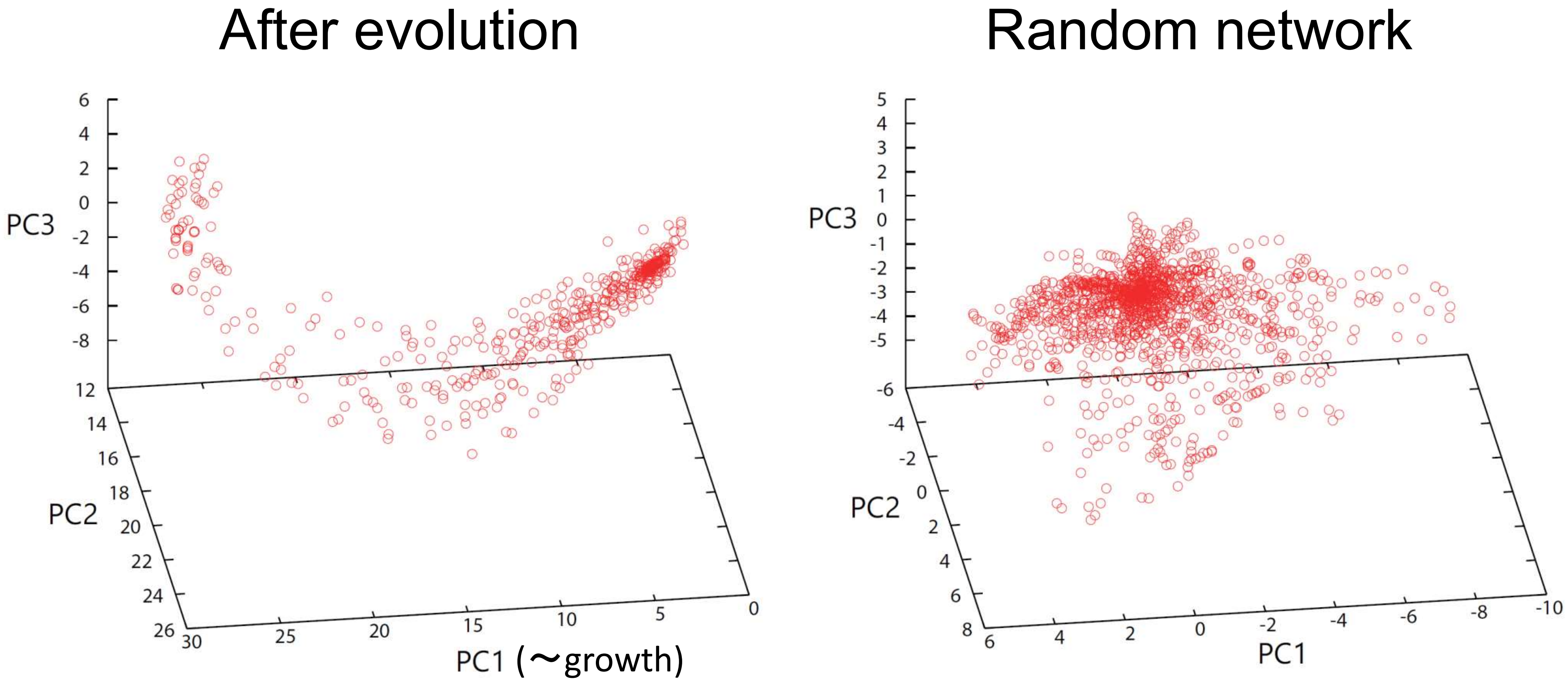
δμ（m）/δμ（m'）

slope



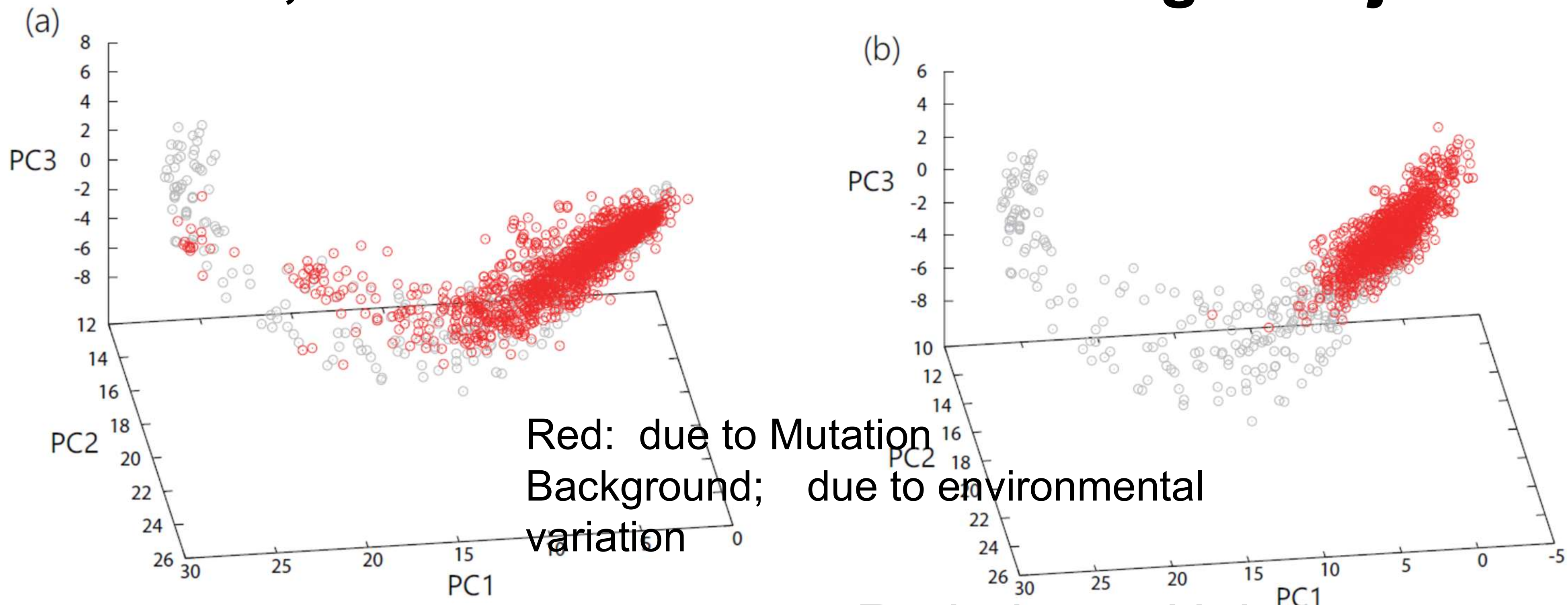
δμ（m）/δμ（m'）

Phenotypic constraint on a low-dimensional space



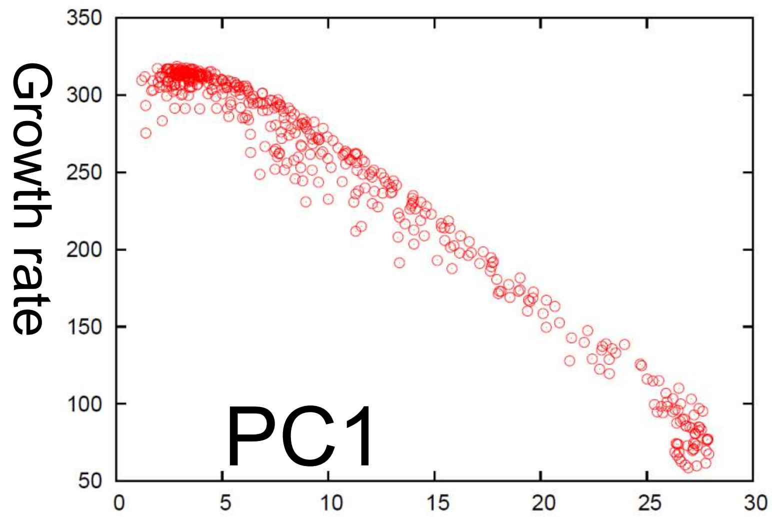
After evolution, the environmental response is constrained on a low-dimensional phenotype space.

Phenotypic change due to environmental variation, mutation, noise are constrained along a major axis



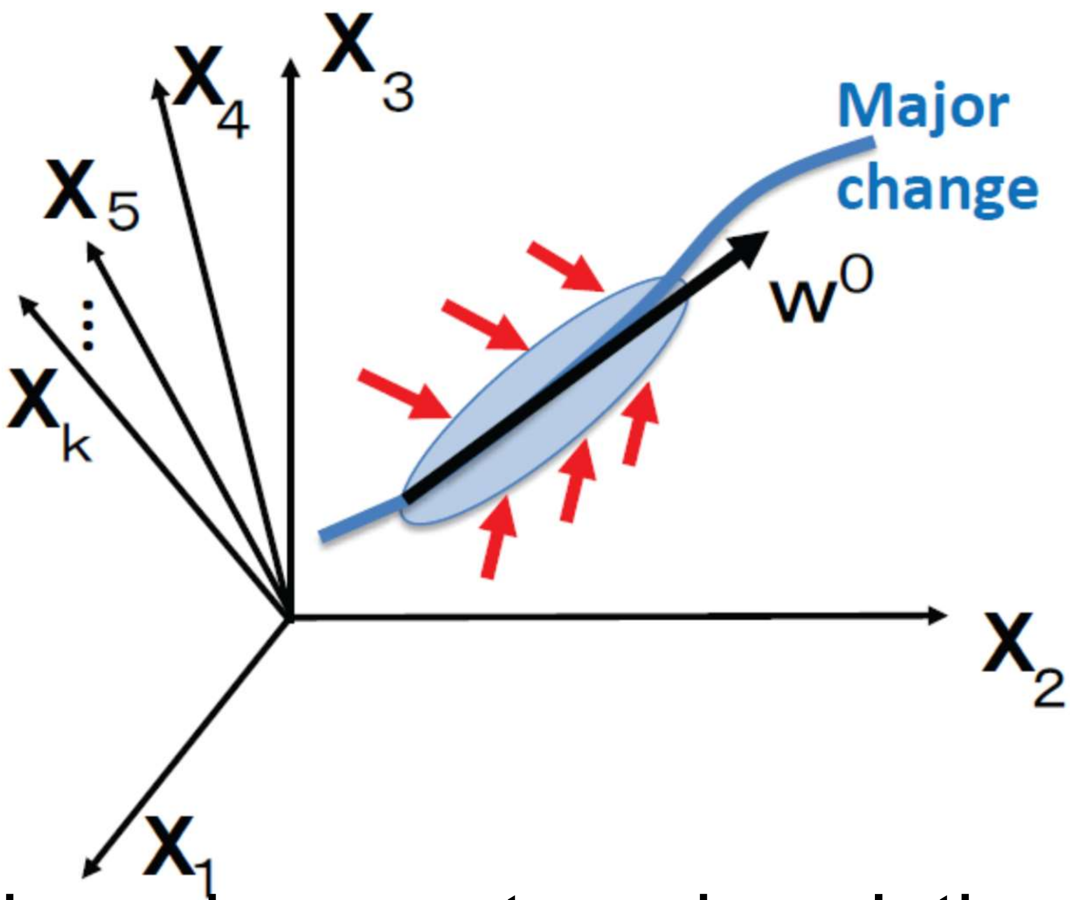
Red: due to Mutation
Background; due to environmental variation

Red: due to Noise
Background; due to environmental variation



?Phenotypic change occurs along a common slow-manifold

Formation of Dominant Mode Along Major Axis



Robust to perturbations – strong attraction from most directions ……

except one direction along which evolution progresses

(Both environment- and evolution- induced) changes in high-dimensional phenotype space are constrained along low-dimensional slow-manifold

Furusawa, KK, Phys.Rev E 2018; KK, Furusawa, Ann Rev Biophys 2018

Formulation and Consequence of Hypthesis

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

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

$$\delta \mathbf{X} = \mathbf{L}(\delta\mu \mathbf{I} - \gamma \delta E)$$

- **γ(E): susceptibility to environment change**

Only the smallest eigenvalue in J (or largest in L=1/J) contributes $|\lambda^i| >> |\lambda^0| \sim 0$

Most changes occur along such slow manifold

$$\delta \mathbf{X} = \lambda^0 \mathbf{w_0}(\delta\mu(\mathbf{v_0}\cdot\mathbf{I}) - (\mathbf{v_0}\cdot\gamma)\delta E).$$

Projection to this manifold w_0

w^0 (v^0) right(left) eigenvector for the smallest eigenvalue, **i.e.,** Projection to this slow manifold

$$\frac{\delta \mathbf{X}(\mathbf{E})}{\delta \mathbf{X}(\mathbf{E}')} = \frac{\delta\mu(E) - (\mathbf{v_0}\cdot\gamma(\mathbf{E}))\delta E/(\mathbf{v_0}\cdot\mathbf{I})}{\delta\mu(E') - (\mathbf{v_0}\cdot\gamma(\mathbf{E}')\delta E'/(\mathbf{v_0}\cdot\mathbf{I})}$$

$\gamma \cdot v_0$ small

Consequence of Slow-Manifold Hypothesis (cont'd)

→Slow manifold is roughly orthogonal to **γ**

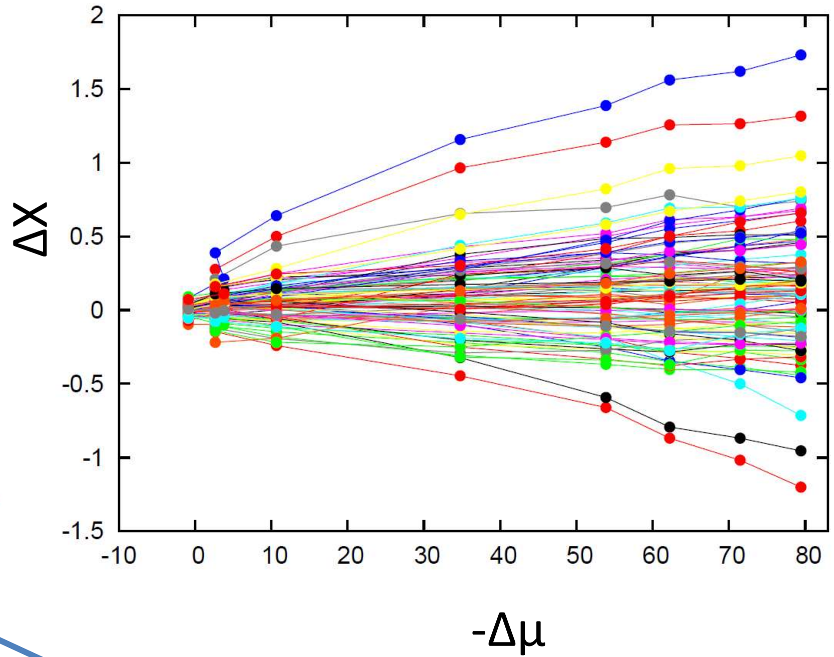
$\boldsymbol{\gamma} \cdot \mathbf{v}_0 \sim 0$

→ $$\delta\hat{\mathbf{X}} = \lambda^0 \delta\mu \mathbf{w}^0$$

Or, from the linear approximation

$$\delta E = \delta\mu / \alpha(E)$$

$$\frac{\delta \mathbf{X}(\mathbf{E})}{\delta \mathbf{X}(\mathbf{E}')} = \frac{\delta\mu(E)}{\delta\mu(E')} \frac{(1 - (\mathbf{v_0} \cdot \gamma(\mathbf{E}) / (\alpha \mathbf{v_0} \cdot \mathbf{I}))}{(1 - (\mathbf{v_0} \cdot \gamma(\mathbf{E}') / (\alpha \mathbf{v_0} \cdot \mathbf{I}))}$$

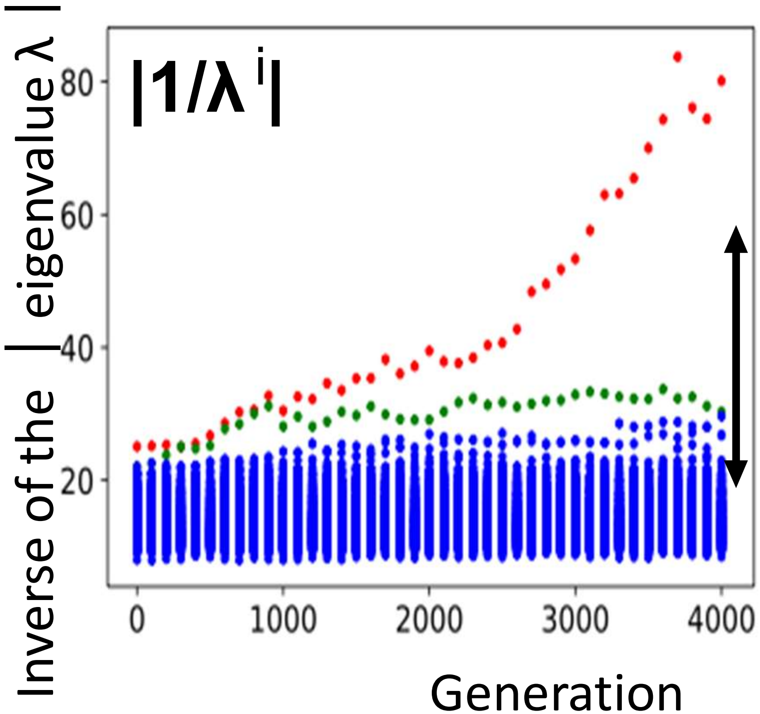


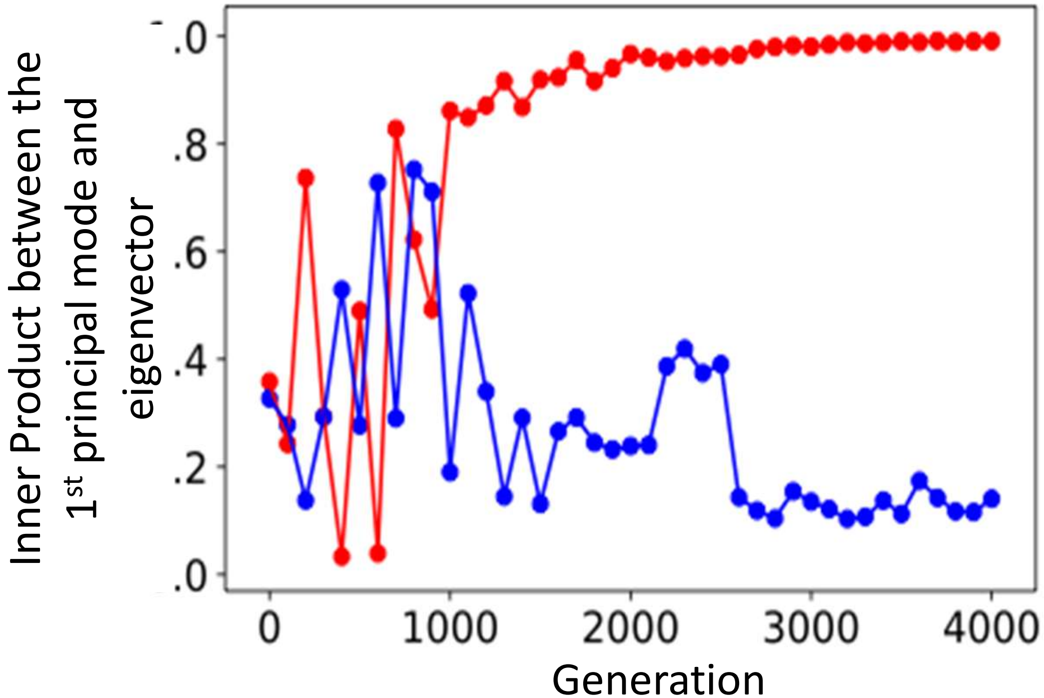
Correction in proportion coefficient

Separation of slowest mode in catalytic reaction net model

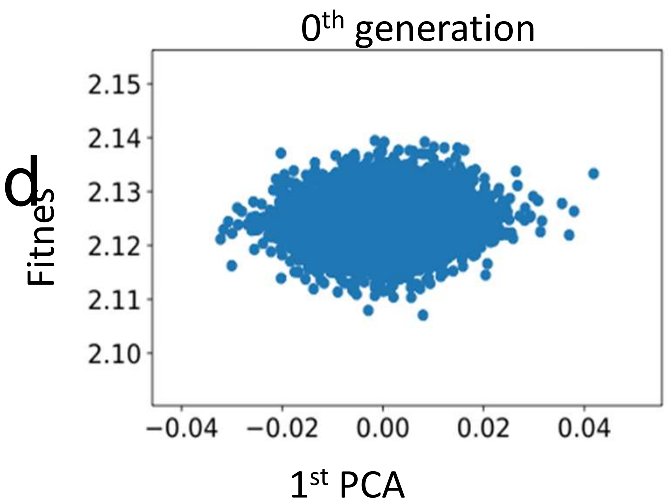
Eigenvalues of $J_{ij} = (\partial \dot{X}_i / \partial X_j)_{\boldsymbol{X}_i = \boldsymbol{X}_i^*}$

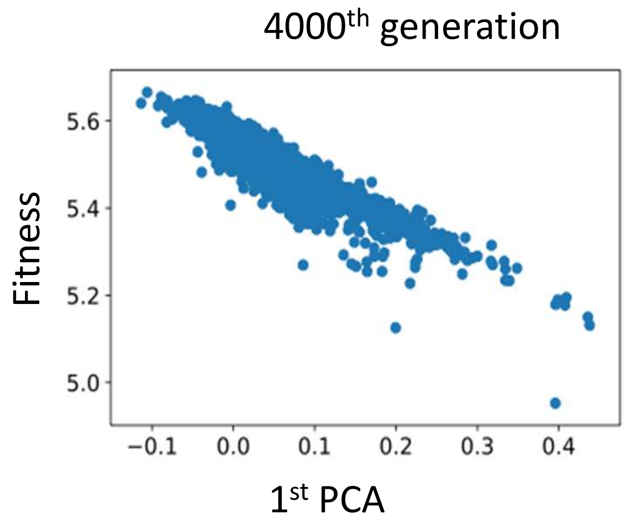
Sato,KK PhysRevR 2020



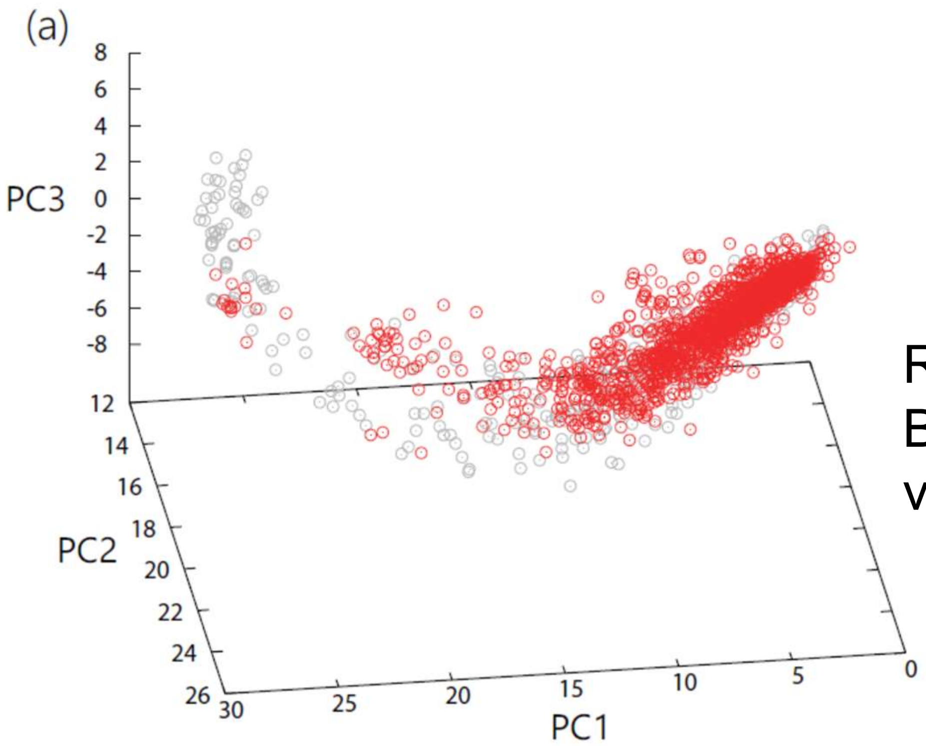


The directions of slowest mode and the fitness are aligned after evolution





→ Evolution -- Recall: Phenotypic change due to environmental variation, mutation, noise are constrained along the same major axis



Red: due to Mutation
Background; due to environmental variation

Phenotypic changes by evolution and environmental changes are along a common dominant mode

Consequence of Hypothesis → Correlation between Environment (**E**) vs Evolutionary (genetic) (**G**) Changes

$$\mathbf{J}\delta\mathbf{X} + \gamma(\mathbf{E})\delta E + \gamma(\mathbf{G})\delta G = \delta\mu(E).$$

Again, assume that

most changes occur along such slow manifold

Project to this slow manifold →

$$\boxed{\delta X_i(G)/\delta X_i(E) = \delta\mu(G)/\delta\mu(E)}$$

using $\gamma \cdot v_0 \sim 0$

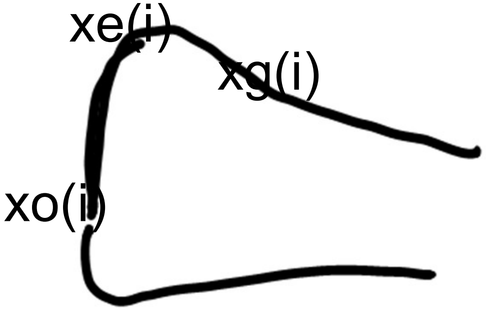
(Genetic) evolution under the environmental condition

→recover growth-- $|\delta\mu(E)| > |\delta\mu(G)|$

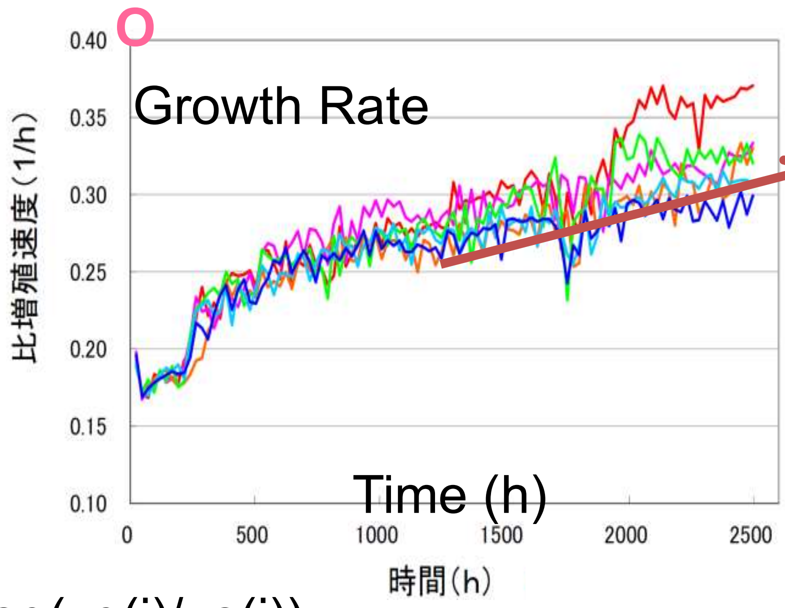
$\delta X_i(G)/\delta X_i(E) = \delta\mu(G)/\delta\mu(E) < 1$

→ All the expression levels tend to return the original level by evolution **Le Chatelier Principle?**

Evolution Experiment of E Coli to adapt in stressed (ethanol) condition



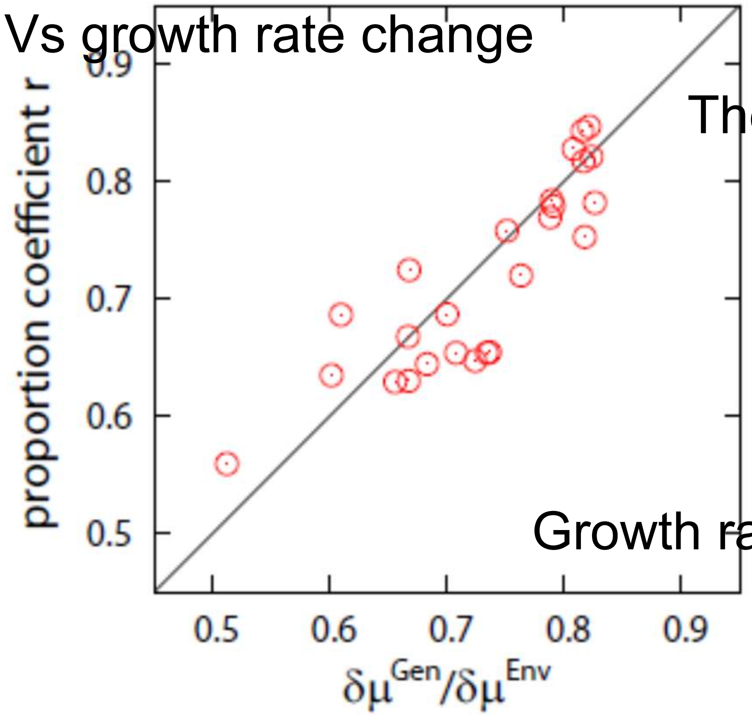
Furusawa'sGroup



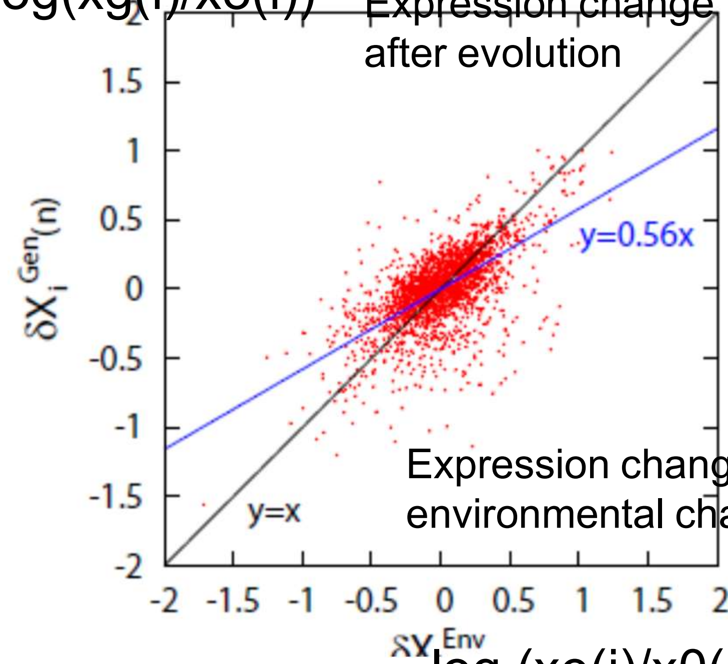
〜1000 generations

Furusawa,KK Interface,2015

Slope in expression change Vs growth rate change



log(xg(i)/xo(i))



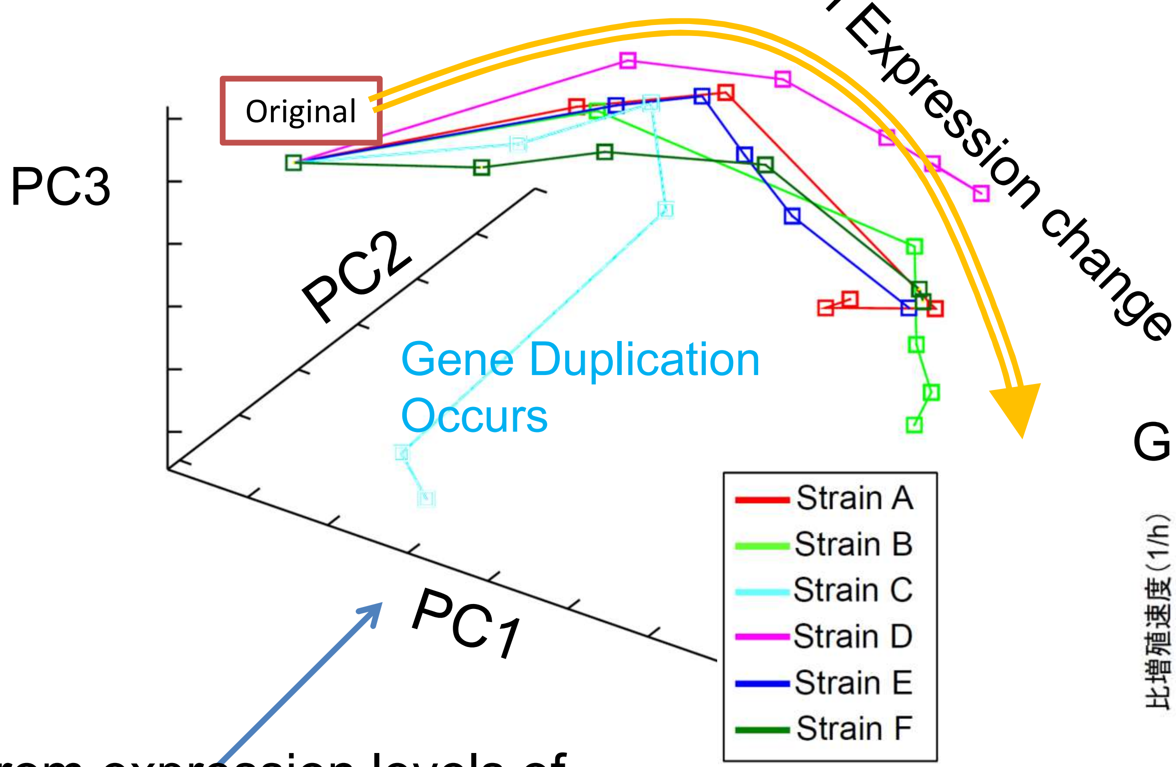
$0 < \delta X_i(E,G) / \delta X_i(E) < 1$

return to original expression pattern

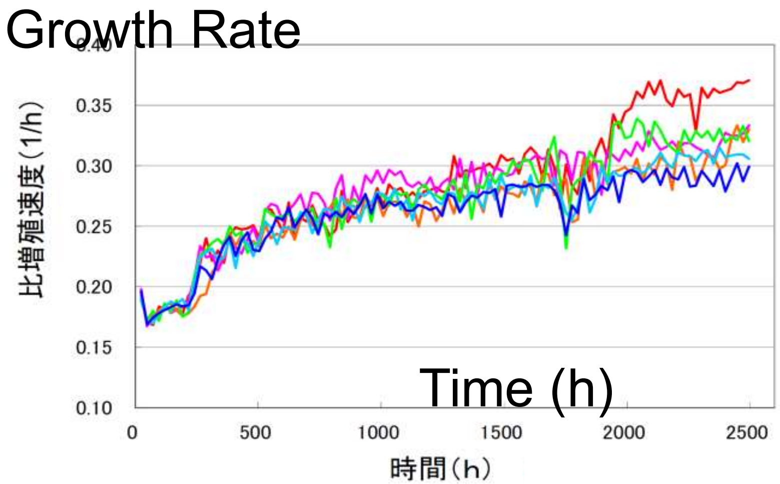
（Le Chatelier principle)

Deterministic phenotypic evolution constrained in low-dim space

Horinouchi,..,Furusawa,
BMC evol Biol 2015



Replaying the tape of evolution, same phenotypic path (not genetic) arises!

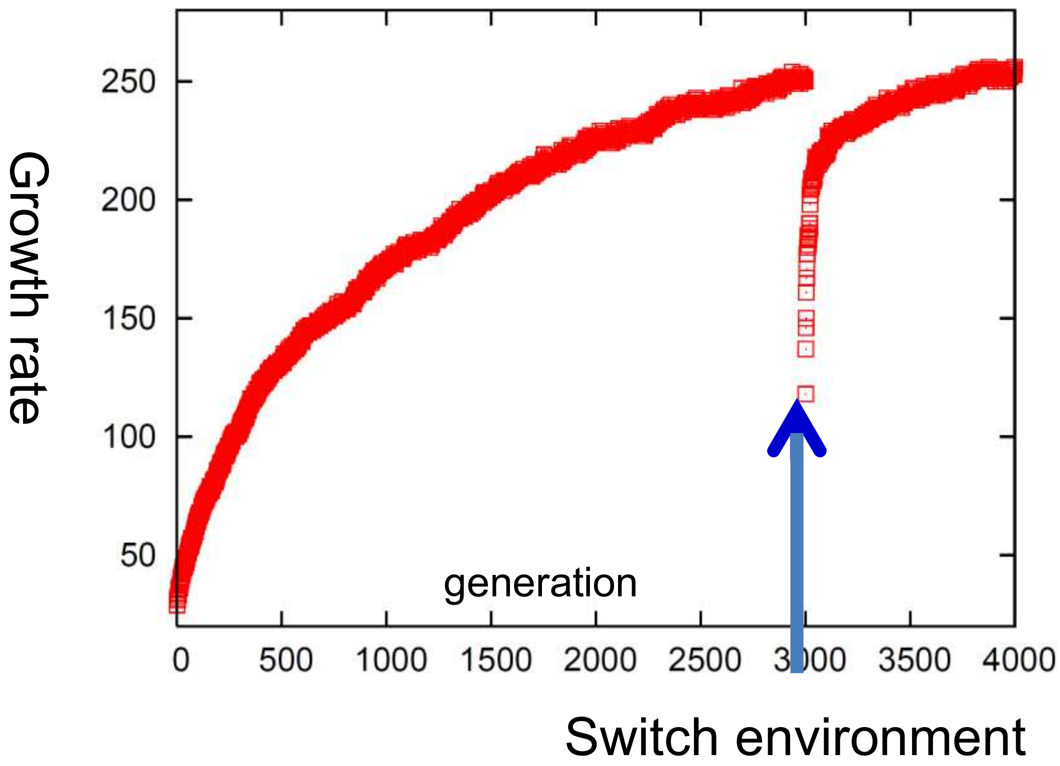


From expression levels of ~4000 genes:

Mutation sites are different by strains.   But..
Common trends in phenotypic space (low-dim structure)
PC1 is highly correlated with the growth rate

Evolution of Catalytic reaction net model by switching environment (nutrient concentratyion) and check evol-env response

Mutate network and select those with higher growth –evo



Recovery of growth rate by adaptive evolution to new environment

$\delta X_i(G)/\delta X_i(E)=\delta\mu(G)/\delta\mu(E) < 1$ (Across all components)

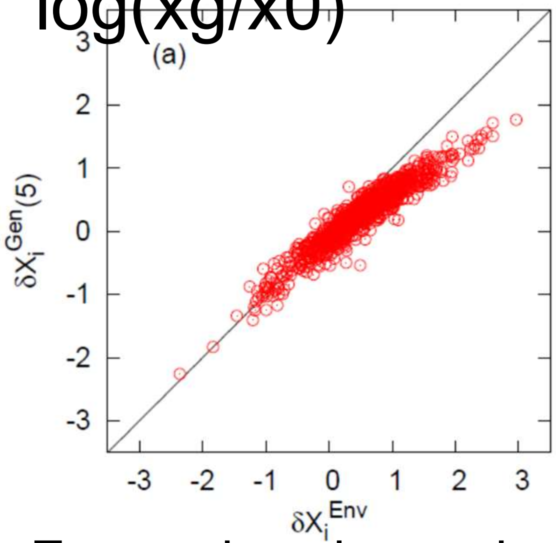
(1) Response by genetic change tends to cancel the change by environment

(2) The two responses are proportional over all components

Expression Change by evolution

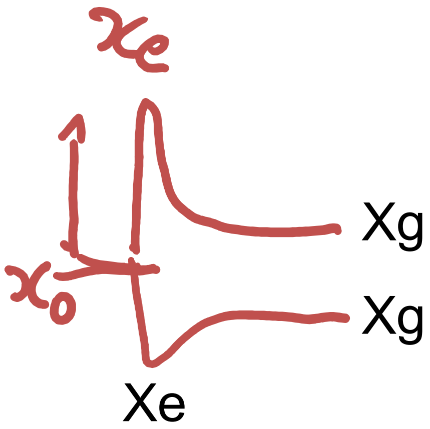
5-th generation

log(xg/x0)

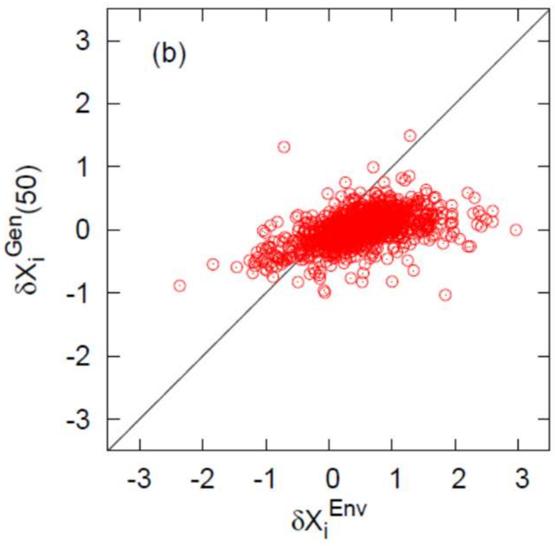


Expression change by env

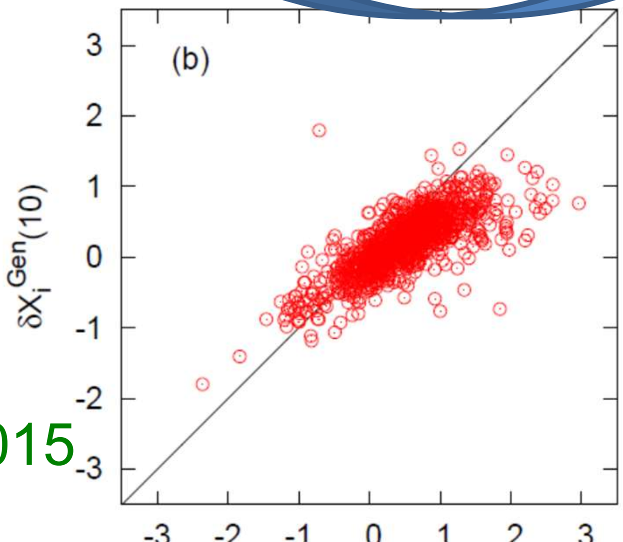
log(xe/x0)



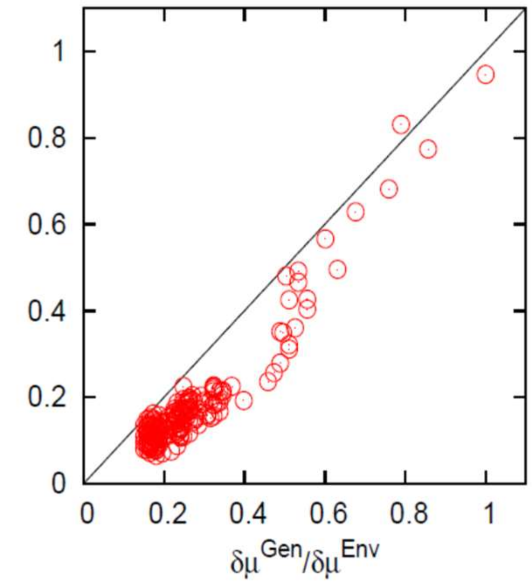
100 th generation



2oth generation



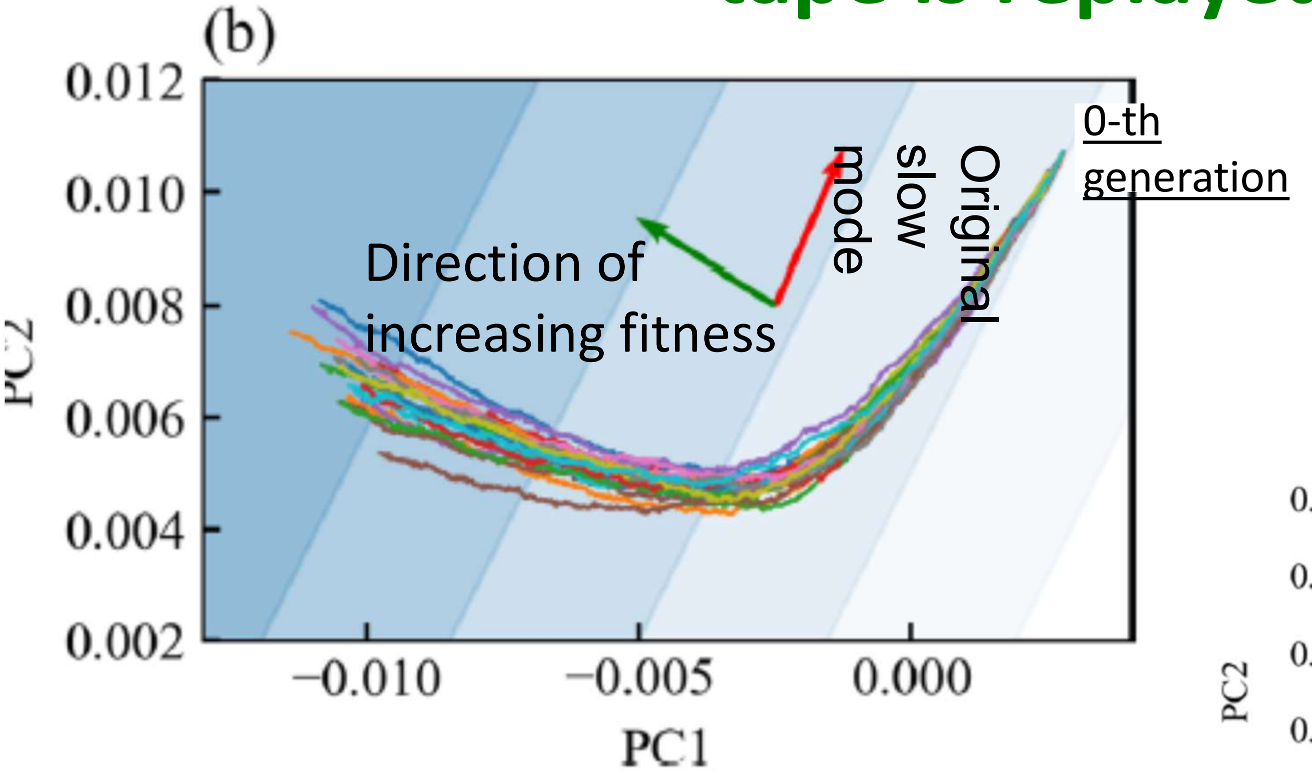
Slope in δX



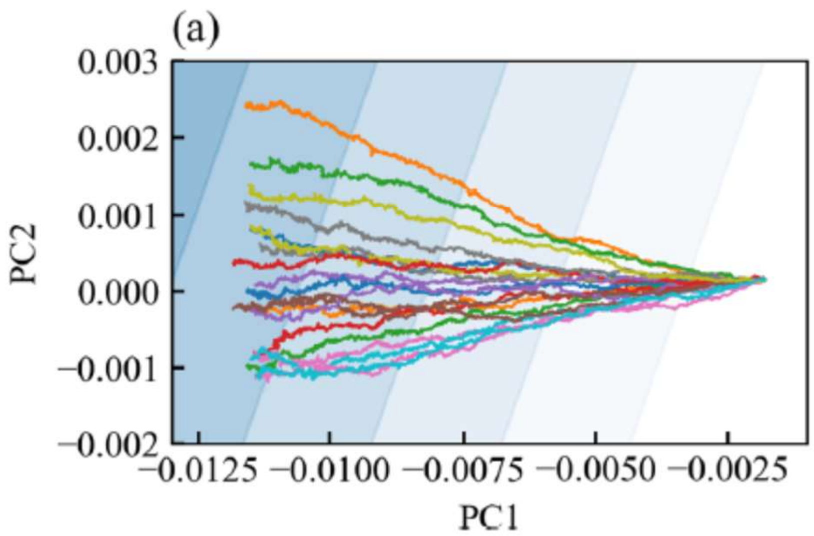
—Δμ bo by env to by evol

Furusawa,KK, Interface 2015

Evolution to novel environment -- the already evolved dominant mode is adopted to adapt to new environment→ **Same phenotypic path when the tape is replayed**.



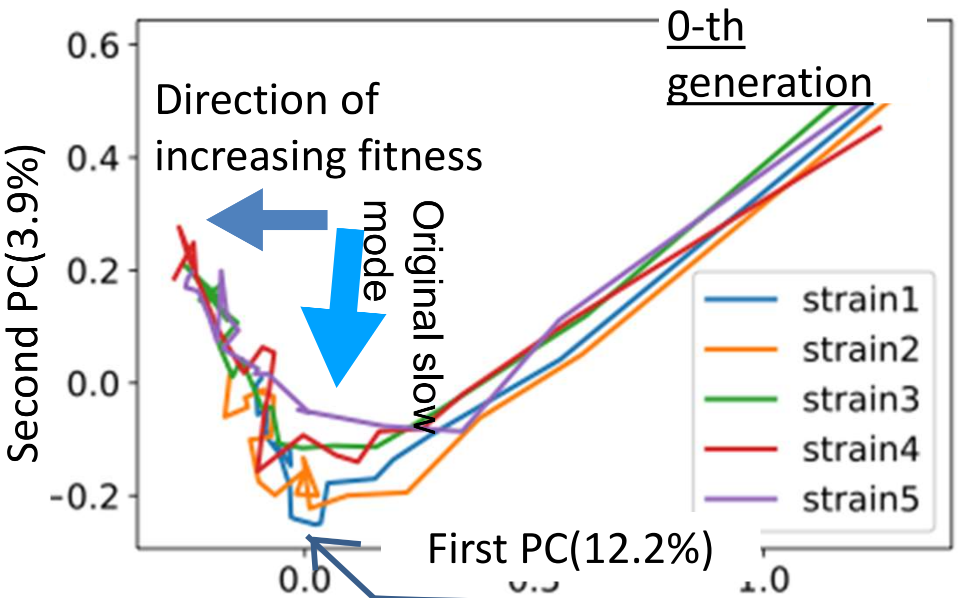
Cf. When started from non-adapted case (same random network)

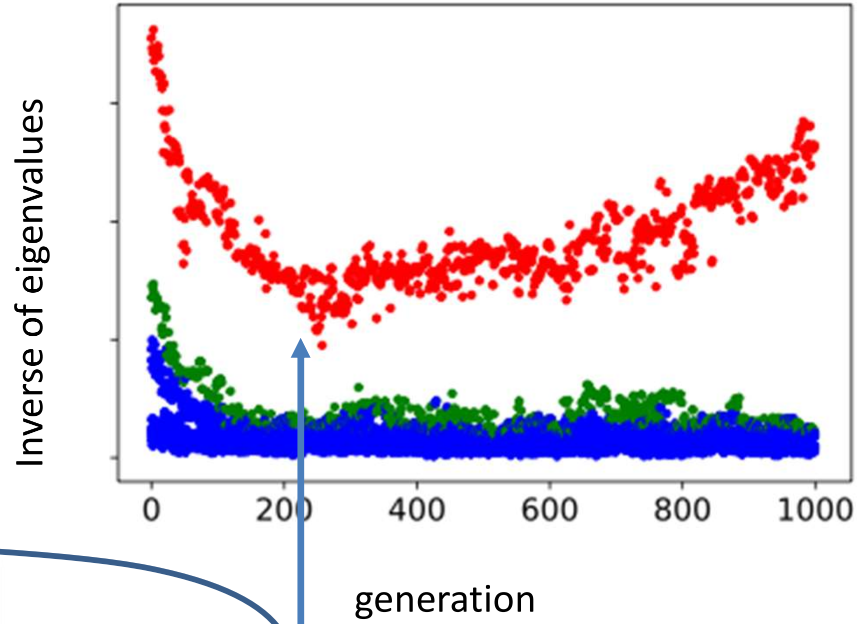


Different color : different strains with different genetic change

Sato, KK, PhysRevRes2020

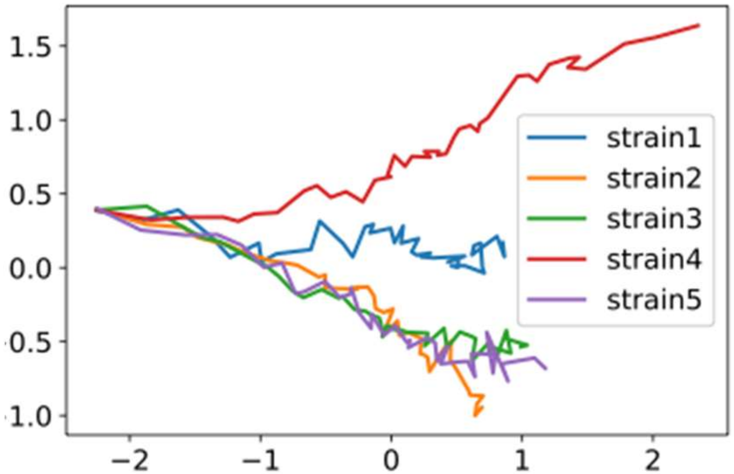
- Evolution to novel environment -- the already evolved dominant mode is adopted to adapt to new environment→ **Same phenotypic path when the tape is replayed**.





Slow mode is adjusted to novel fitness

Cf. When started from non-adapted case (same random network)

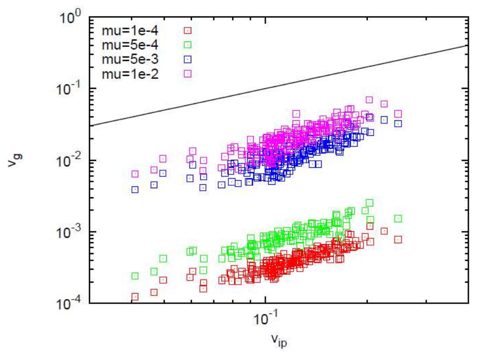


Sato, KK, PhysRevRes2020

So far response relationship: fluctuation & response are two sides of the same coin (←Einstein)

Fluctuation

Variance by gene change Vg



~proportional

Variance by noise Vip

←classic Fisher Theorem→

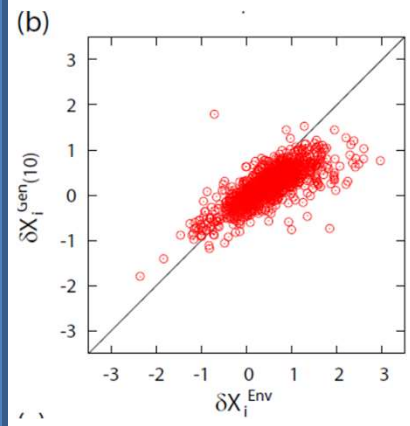
Evolutionary Fluctuation-Response (2003)

← Proportion →

Response by evolution

Genetic change

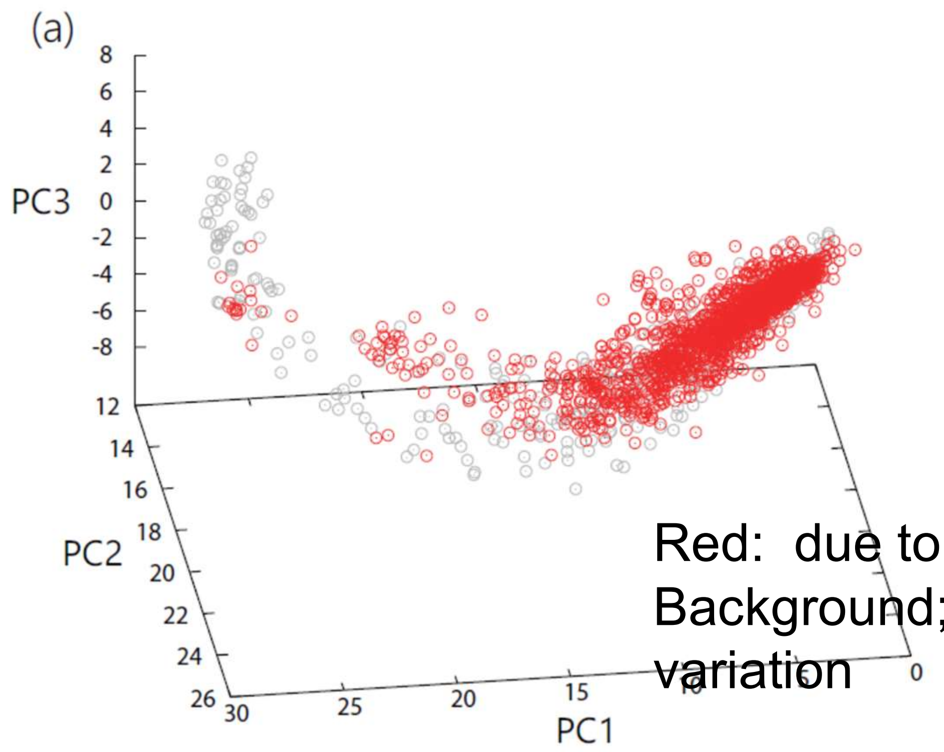
~proportional



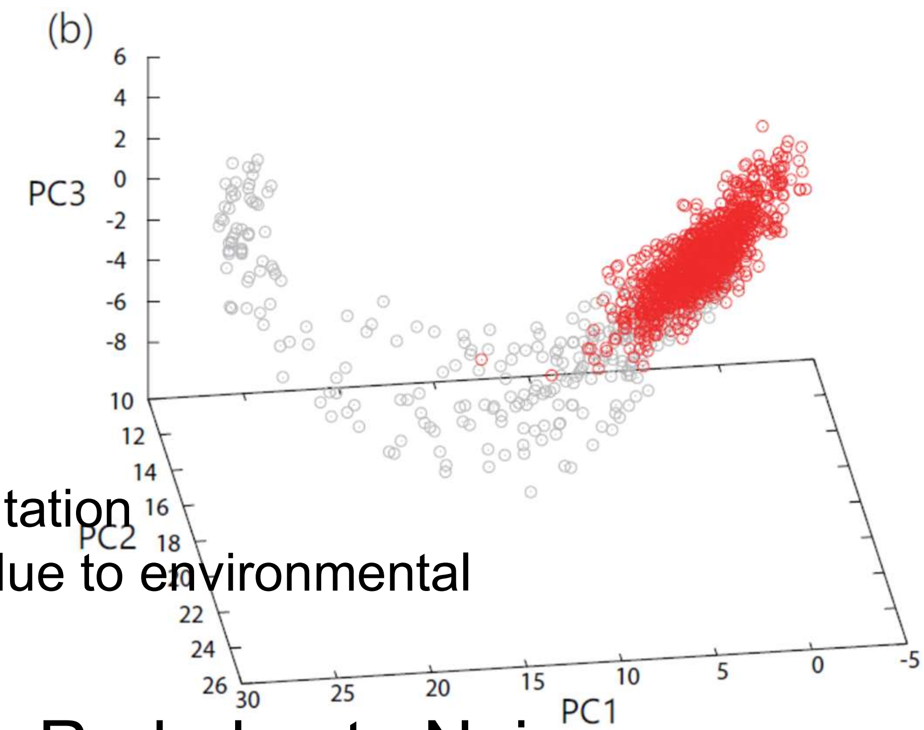
Response by environ-ment

Non-genetic change (noise, environ-ment)

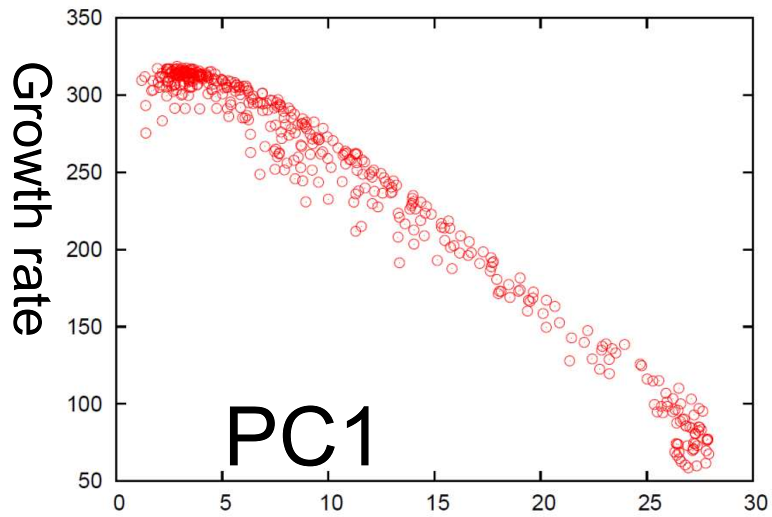
Recall…



Red: due to Mutation
Background; due to environmental variation



Red: due to Noise
Background; due to environmental variation



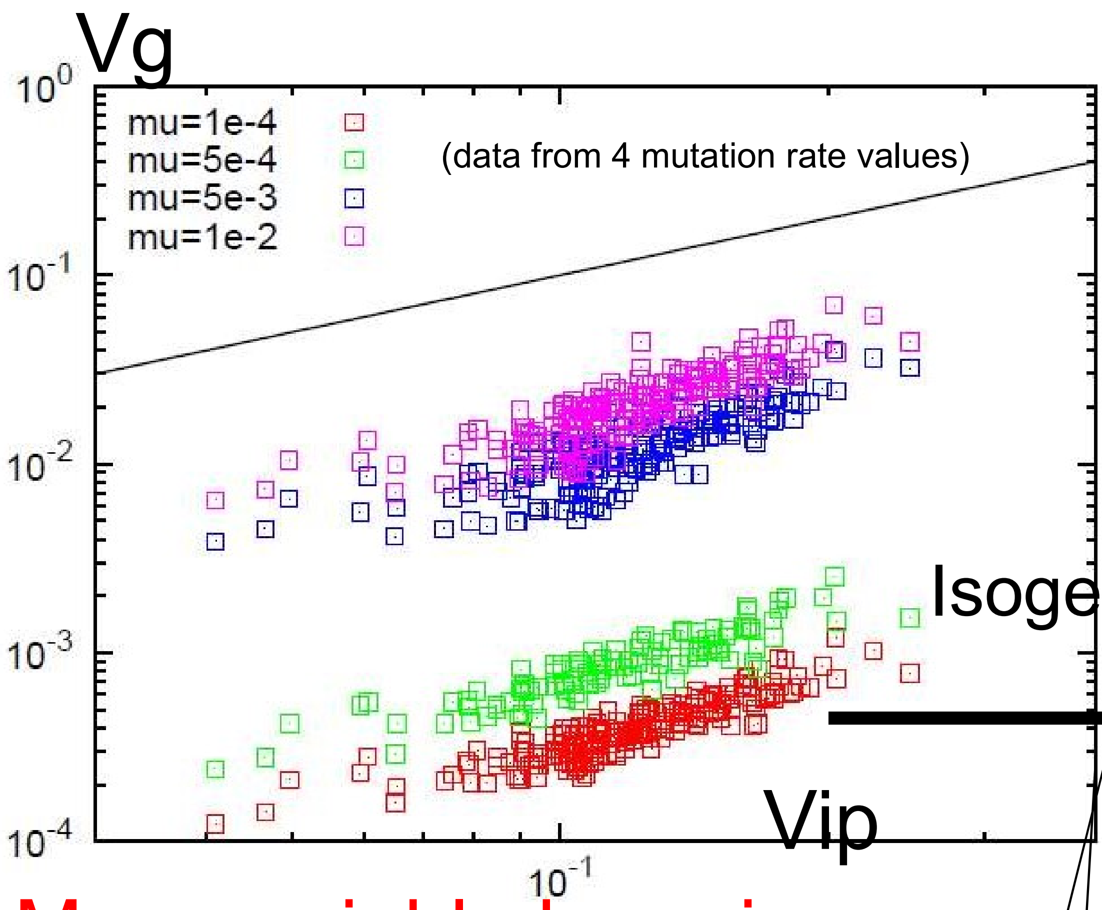
?Phenotypic change occurs along a common slow-manifold

Furusawa, kk Interface 2015

Vip-Vg relationship across traits (phenotypes)

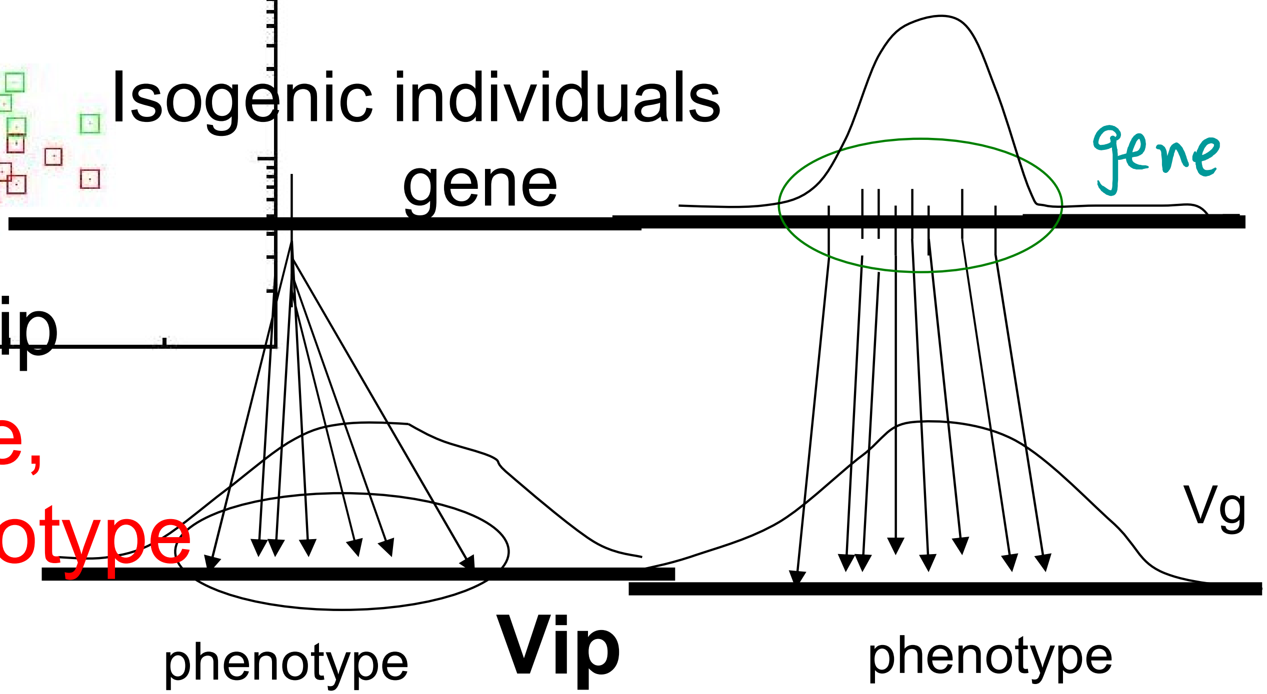
Vg(i)：Vatiance of X(i) due to genetic mutation

Vip(i)：Variance of X(i) due to noise in dynamics



Vip=Vg

Vip(i)∝Vg(i)∝evol.speed
over all traites i



More variable by noise,
more evolvable: Phenotype
evolution predictable

Vg-Vip proportionality is explained by the slow manifold Hypothesis

Evolution occurs along this dominamt manifold $\mathbf{w}^0$

$$V_{ip}(i) = (\mathbf{w}_i^0)^2 < \delta X^2 >_{noise}$$

$$V_g(i) = (\mathbf{w}_i^0)^2 < \delta X^2 >_{mutation}.$$

→ Vg(i)/Vip(i) = independent of i

(here we do not need the growth-rate constraint, only slow-manifold constraint is needed)

Vg-Vip relationship ← Changes both by (environmental) noise and (genetic) mutations are constrained along the direction

Need further studies to establish the present theory

(i) Further Confirmation by Experiments

(ii) Confirmation by Models : Universality?

Catalytic Reaction Net-Cell Model ☑

Gene regulation Net Model (Sato, KK in prep) ☑

Spin-glass Models (Sakata KK., PRL 2020) ☑

evolve spin Hamiltonian JijSiSj to achieve certain configuration

dimensional reduction at replica symmetric phase

Protein Model/Data (Tang KK., PRL2021) ☑

correlation in structure dynamics & evolutionary dim reduction

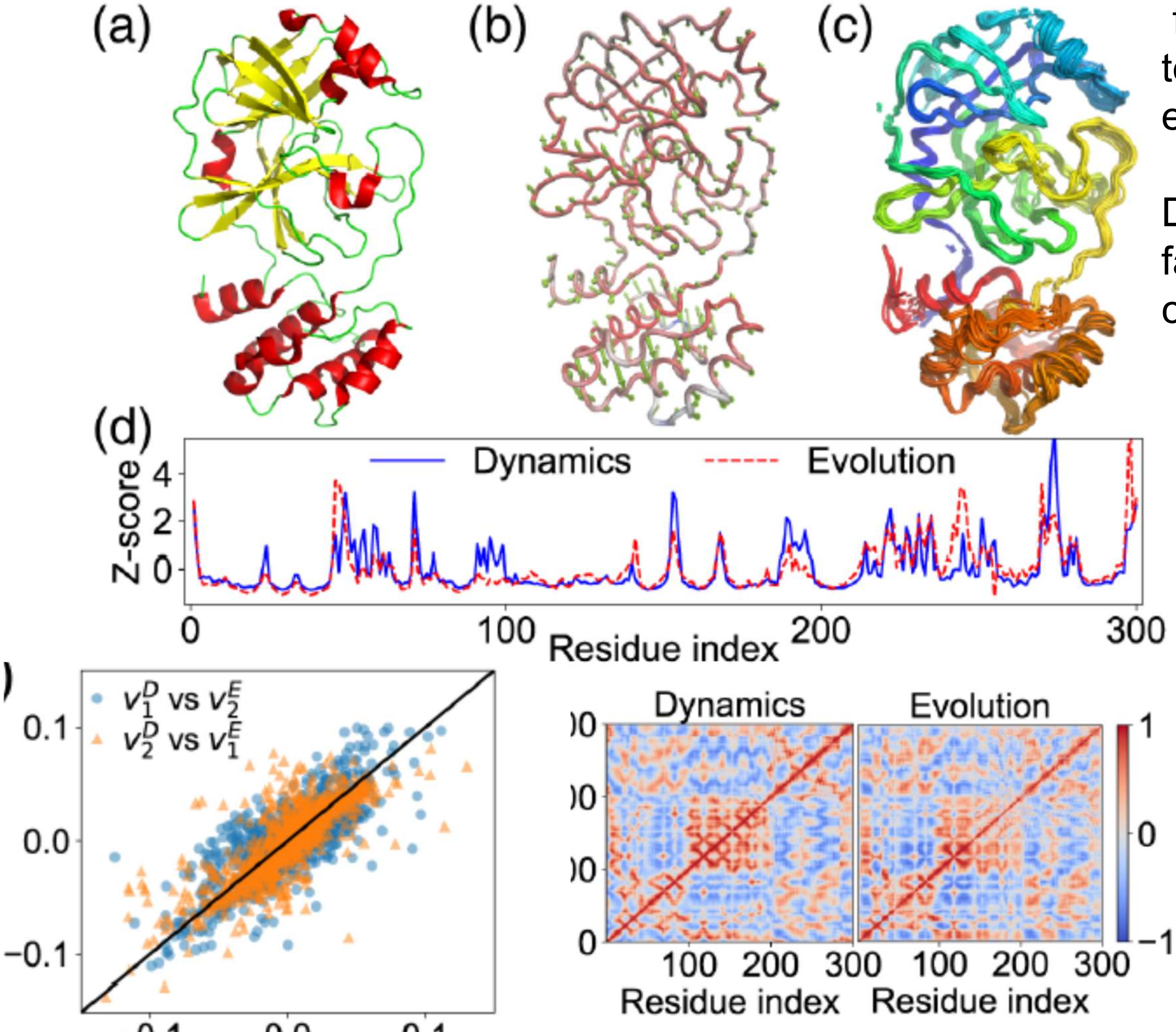
(iii) Theory for dimensional reduction? –1 or few dim?

outliers in eigenvalues – separation of slow modes,

Renormalization Group???

Projection to Collective Modes?

☑Protein; Change in Native structure by noise & by evolution, highly correlated and low-dimensional



fluctuation according to structural data+ elastic net model vs

Difference within family (mutational change)

Changes are lo-dim, and correlated

Tang,KK PRL2021

Spin-Statistical Model Sakata,KK,PRL 2020

Phenotype=Spin config.Si Genotype—Interaction Jij

Hamiltonian H=-ΣJijSiSj

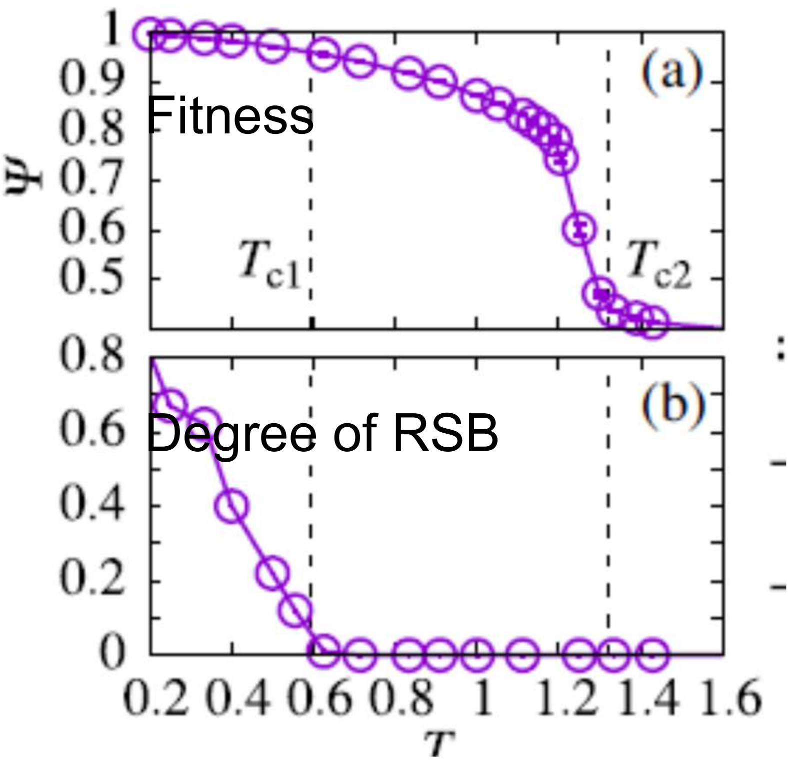
Fitness align target spins; environment– external field

$$\psi(\boldsymbol{J}) = \overline{|m_{\mathcal{T}}|}, \qquad m_{\mathcal{T}} = \frac{1}{N_{\mathcal{T}}} \sum_{i \in \mathcal{T}} S_i,$$

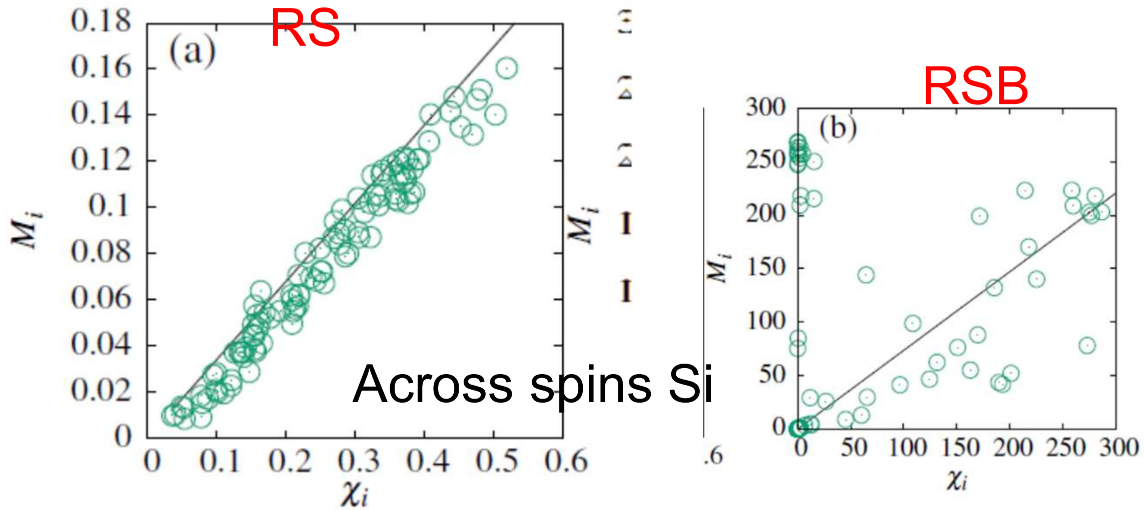
1) Robust fitted state at Replica Symmetric phase

2) RSB → loss of robustness

(cf Sakata,Hukushima,KK PRL 2009)



Correlation in Responses to ext field and to mutation to Jij



Fraction of matrices J in which the RR

Evo-Devo Congruence
Numerical Evolution of development

Cells in 1-dim line

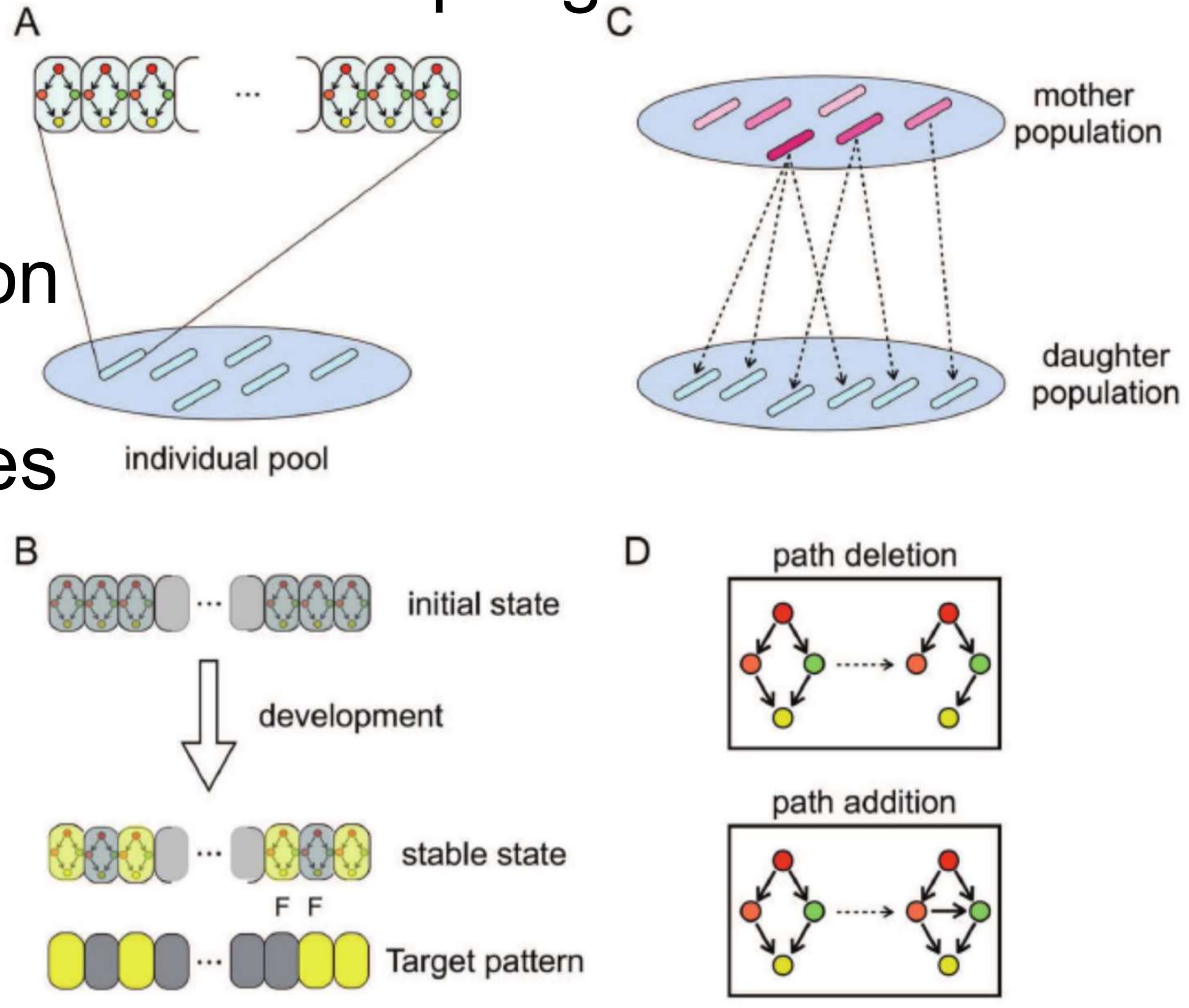
Each cell has protein expression dynamics by GRN

External morphogen gradient for input genes

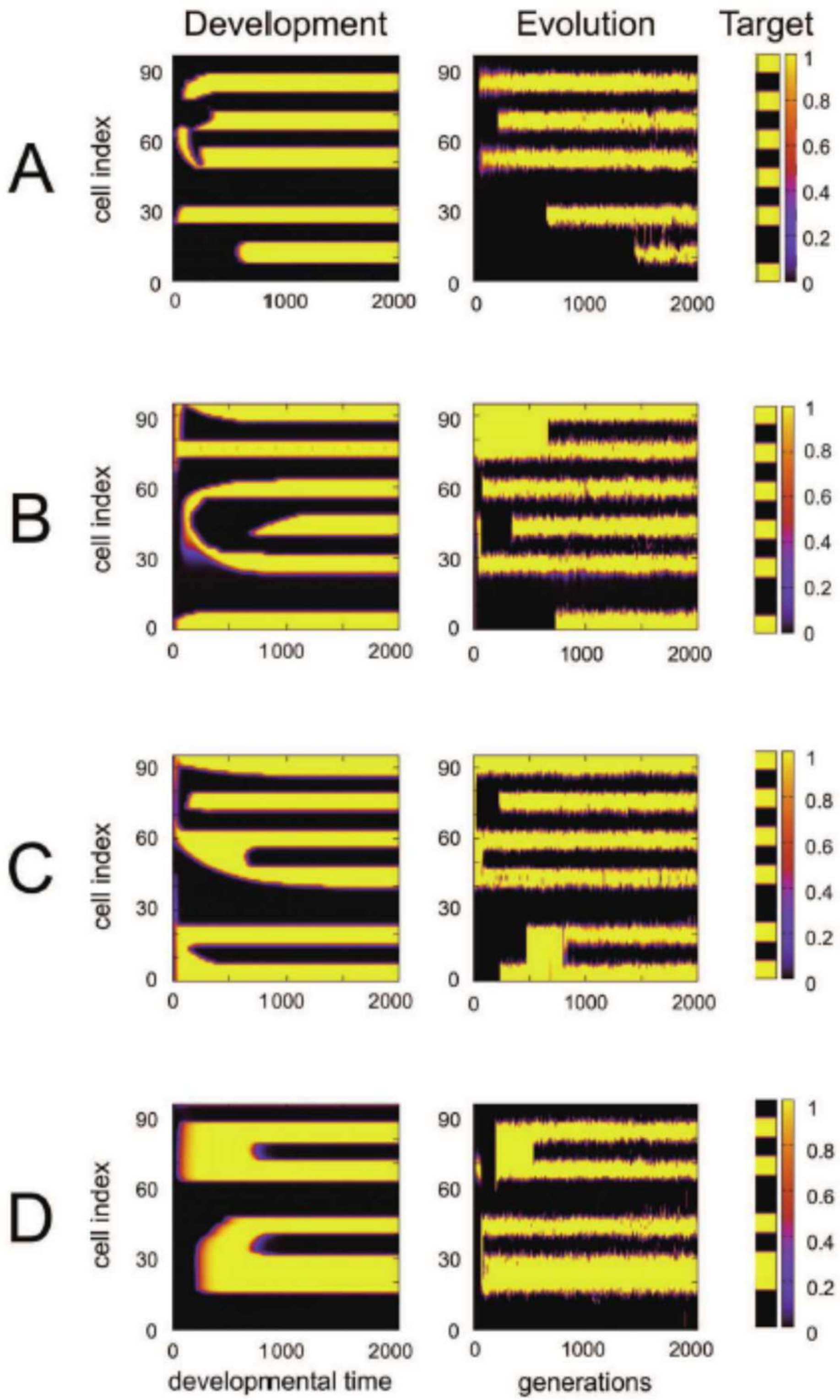
diffusion of proteins

Evolve GRN by mutation

Fitness: Given target pattern for output genes



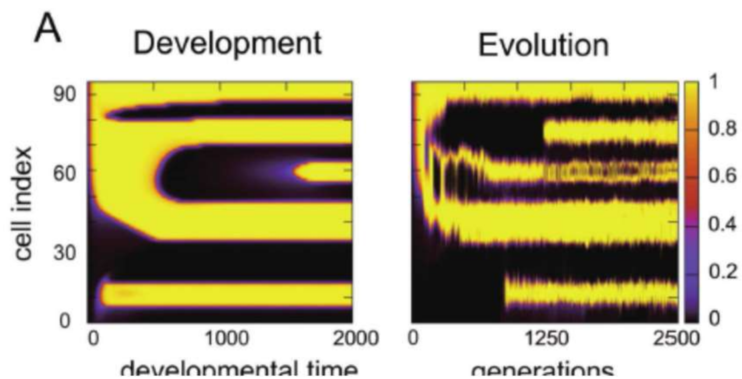
Kohsokabe & KK J Exp Zoology B (2016,2021)



Congruence between development and evolution (cf, Haeckel,recapitulation)

For most (95%) examples, good correspondence

Rare exception



Messages

- (Cellular) Phenotypes are high-dimensional, but their adaptive changes are drastically restricted in a low-dimensional space
- Slow modes evolve and fascillitate evolution

← Result of steady-growth and evolutionary robustness (to noise and to genetic changes)

- Phenotypic evolution is rather deterministic even though genetic changes can be stochastic

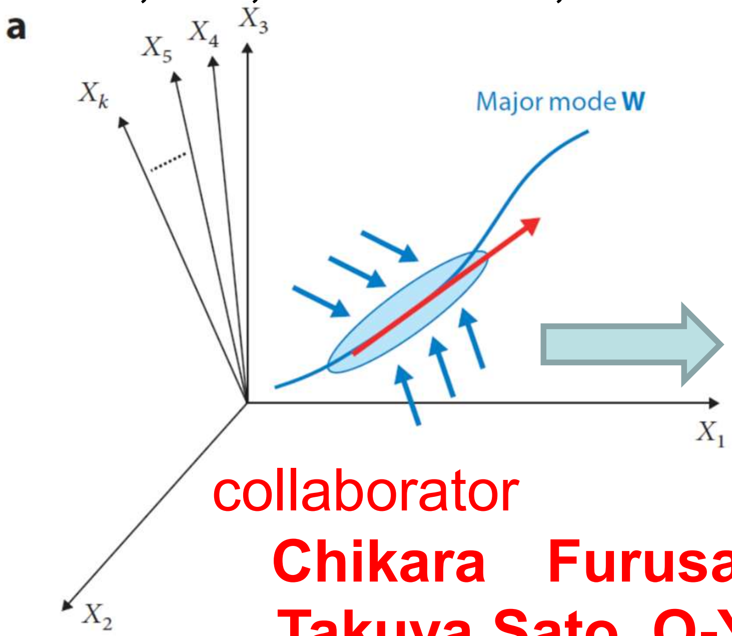
(replaying the tape, phenotypically same path)

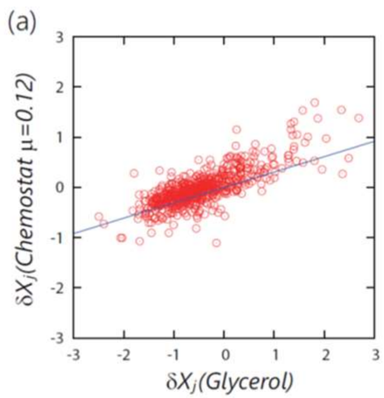
← Phenotypic evolvability correlated by short-term dynamics and fluctuation

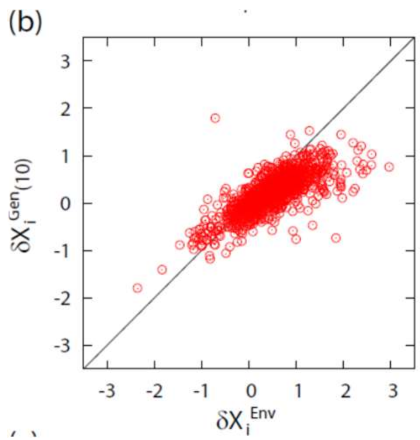
Summary

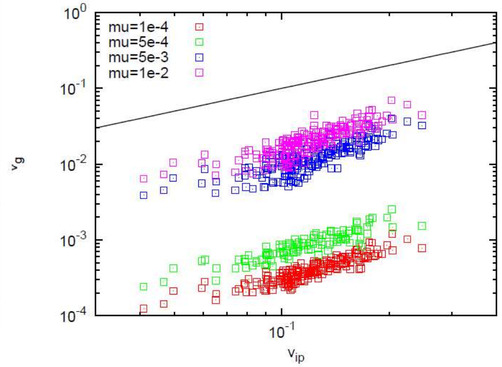
Low-dimensional structure formed from high-dimensional phenotypic space ← robustness

(Furusawa, KK, Phys Rev E, 2018; KK, Furusawa, Ann Rev Biophys 2018; Sato, KK, PRR 2020; Sakata, KK, PRL 2020, Tang KK PRL 2021)









collaborator

Chikara Furusawa ;
Takuya Sato, Q-Y Tang
A.Sakata

Universal law for adaptation

(KK Furusawa Yomo PRX2015)

Evolutionary LeChatelier Principle

(Furusawa KK Interface 2015)

Vg-Vip Law (→ direction in phenotypic evolution)