Function and Robustness of Gene Regulatory Network: Toward the Landscape Picture of Evolution

Macoto Kikuchi (Collaborators: S. Nagata and T. Kaneko)

大阪大学サイバーメディアセンター

2019/3/7

Contents

Motivation

Characteristic properties of "evolved thing" are function and robustness

A.Wagner: "Robustness and Evolvability in Living Systems" (2005)

- Intuitively, highly optimized system may be fragile.
  - Evolution is not simply an optimization process?

Robustness

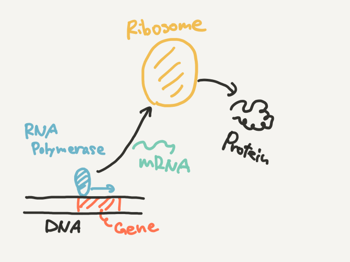
- Robustness against perturbation
  - Stability in development and differentiation: Canalization (Waddington)
    - Epigenetic landscape
  - Protein folding: Anfinsen's dogma, Funnel picture (Go, Wolyness)
    - Energy landscape
- Robustness against mutation
  - Function is not lost by mutation
  - Homologous protein

Prospect

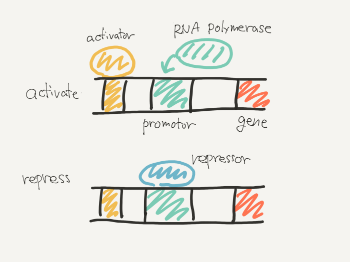
- Landscape picture of evolution
  - Consider evolution landscape, including phenotypes not visited in the course of evolution
  - Evolutional pathway on the landscape

- We consider a toy model of the gene regulatory network
- As the evolved system should be rare, we use the rare event sampling method

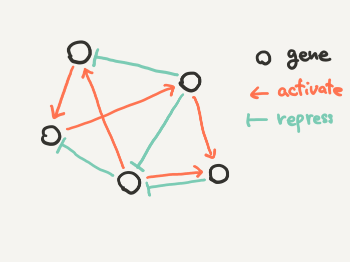
Gene Regulatory Networks (GRN)



Gene expression



Gene regulation



Abstract model of GRN

- A complex network in which the genes mutually regulate by the transcription factors (TF)
  - TFs themselves are proteins made by the gene expressions

Question

- Character of the fitness landscape
- Relation between the cooperative response to outside and the robustness
    - Mutational robustness
    - Robustness against external/internal fluctuation (number fluctuation of TF or other molecules)

Can we see any universal properties, if we classify the randomly generated GRNs by fitness

- Properties that do not depend on the evolutional pathway

Model

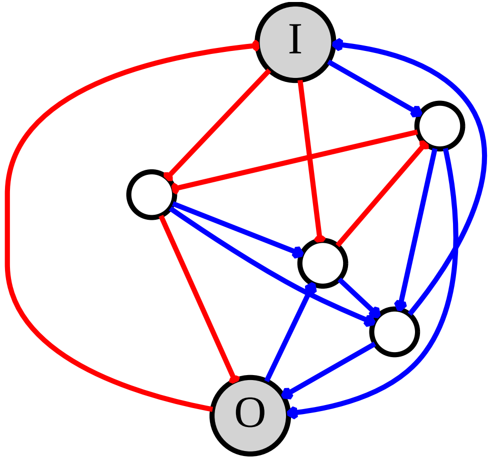
Simple toy model of GRN having one input gene and one output gene

- cf. M. Inoue and K. Kaneko PLOS Compt. Bio. **9**(2013)e1003001

Directed random graph: N nodes、K edges

- Node: Gene
- Edge: Regulatory relation

- Self regulation and mutual regulation are excluded
- The input node is randomly selected from the nodes having paths to all the other nodes
- The output node is selected from the nodes faving paths from all the other nodes (Detail



GRN having one input node and one output node and having no self and mutual regulations

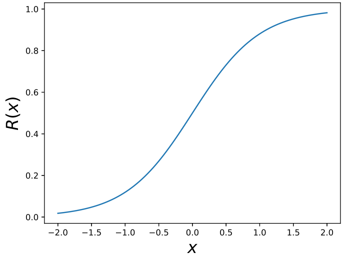
Discrete-time dynamics (Neural-network like)

$$X_i(t+1; I) = R\left(I\delta_{j,1} + \Sigma_j J_{ij} X_j(t; I)\right)$$

$$R(x) = \frac{\tanh x + 1}{2}$$

- cf. A. Wagner: Evolution **50** (1996) 1008

- X_i : Expression of ith gene ($[0,1]$)
- J_{ij} : Regulation of ith gene by jth gene$(0, \pm 1)$
  - $+1$: activation, -1: repression
- I : Input from exterior world ($[0,1]$)
- $R(x)$: Soft response function



Response function

- Spontaneour expression is 0.5 (comparatively large)
- We want to assemble a circuit that can respond sensitively to On and Off of external signal
    - cf. M. Inoue and K. Kaneko:EPL **124** (2018) 38002

Required function

Sensitive response to On-Off change of external signal

- Since the response damps out for sequential circuit, Feed-Forward type regulation is indispensable
- Both activation and repression are required

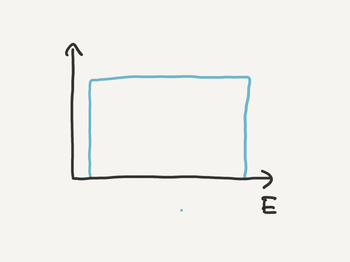
Fitness

- $\bar{X}_i(I)$: Temporal average of the response of ith gene (in the steady state)
- Sensitivity of ith gene: Defference of the response to $I = 0$ and 1

$$S_i = |\bar{X}_i(1) - \bar{X}_i(0)|$$

- The node having the largest sensitivity is defined as the output node
- X_{out}: X of the output node (response of the network)
- Fitness$f \equiv S_{out}$: Sensitivity of the output node

Method (Multicanonical MC)



Ideal energy histgram obtained by the multicanonical MC

- Sampling method that gives a flat distribution of energy
  - Enable us to sample low-energy rare states
  - Enable us to calculate the density of states

- Detailed balance

$$w_{ij}P(E_j) = w_{ji}P(E_i)$$

  - For ordinary Metropolis MC, $P(E) \propto e^{-\beta E}$
- We can use any $P(E)$

$$P(E) \propto e^{-f(E)}$$

and we require

$$e^{-f(E)} \sim \frac{1}{\Omega(E)}$$

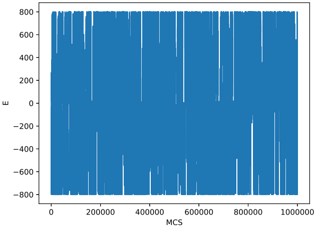
:Weight $f(E)$ is determined through learning process

- Using the obtained energy histgram, we can estimate DOS

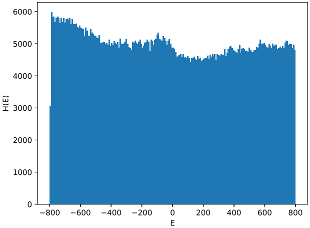
$$\Omega(E) \propto H(E)e^{f(E)}$$

- Divide E into bins
  - Piecewise linear approx. for $f(E)$: Multicanonical
    - B.A. Berg and T. Neuhaus: PRL **68** (1992) 9
  - Constant $f(E)$ in each bin: Entropic sampling
    - J. Lee: PRL **71** (1993) 211
- Wang-Landau method for the learning process
  - used only for the entropic sampling
    - F. Wang and D.P. Landau: PRL **86** (2001) 2050

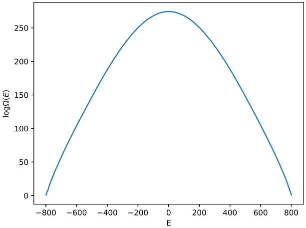
2D Ising Model



Time series of energy



Obtained energy distribution



DOS estimated by multicanonical MC

- Number of the ground state is 2.07 (cf. true value is 2)

Application to non-energetic system

- Eigenvalue distribution of random matrix
  - N. Saito, Y. Iba and K. Hukushima: PRE **82** (2010) 031142
- Search for periodic orbits in a chaotic system
  - A. Kitajima and Y. Iba: Compt. Phys. Comm. **182** (2011) 251
- Stability of a coupled chaotic map
  - N. Saito and M. Kikuchi: New J. Phys. **15** (2013) 053037
- Enumeration of magic squares
  - A. Kitajima and M. Kikuchi: PLOS One **10** (2015) e0125062

The first paper that discuss the evolutionary landscape using multicanonical MC

"Robustness leads close to the edge of chaos in coupled map networks: toward the understanding of biological networks"
N. Saito and M. Kikuchi: New J. Phys. **15** (2013) 053037

Evolution and robustness of a coupled chaotic map (an abstract model for GRN)

Application to GRN

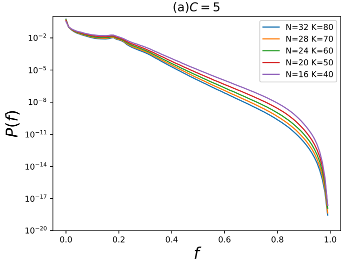
- Sampling that gives the flat distribution of fitness
  - Divide the fitness $(0 \sim 1)$ into 100 bins

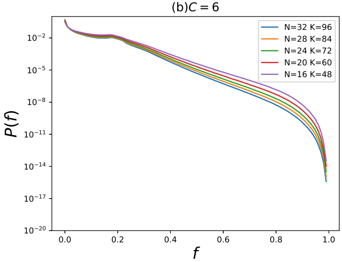
- In principle, we can randomly sample GRNs with several different values of fitness
  - Actually there are correlations between samples
- Microcanonical ensemble within each bin

- $N = 16 \sim 32$
- Average number of edges connected to each node$C \equiv 2N/K = 5, 6$
    - We show results for $C = 5$ mainly

Results 1

Fitness Landscape





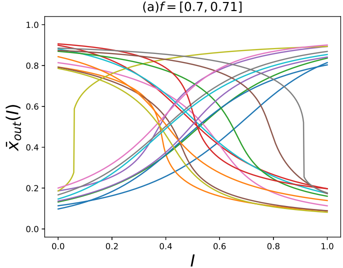
Appearance probability of fitness ($C = 5$)

Appearance probability of fitness ($C = 6$)

- There is a threshold of rareness
  - more than 95% in $f < 0.2$
- GRNs having f larger than the threshold are exponentially rare
- $f > 0.9$ are more than exponentially rare
  - $f > 0.99$: The fittest ensemble

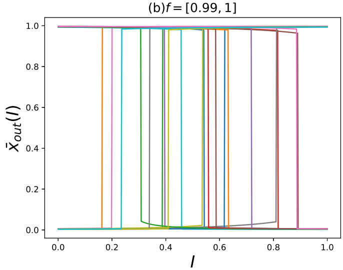
GRNs with high fitness are rare

Response in steady states



$f = [0.7, 0.71]$ (20 samples)

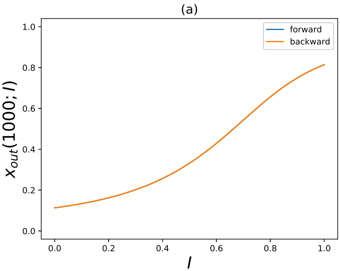
- Steady-state response when the initial condition is $S_i = 0.5$ for all i
  - Smooth response to the input I
  - A single fixed point



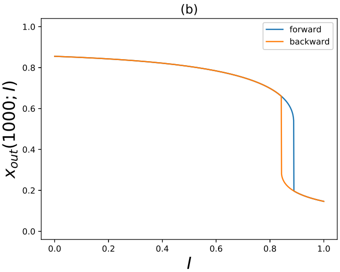
The fittest ensemble (20 samples)

- Step-like response to the input I
  - Response by switching two fixed points
  - Ultrasensitivity

Responses for $f \sim 0.7$

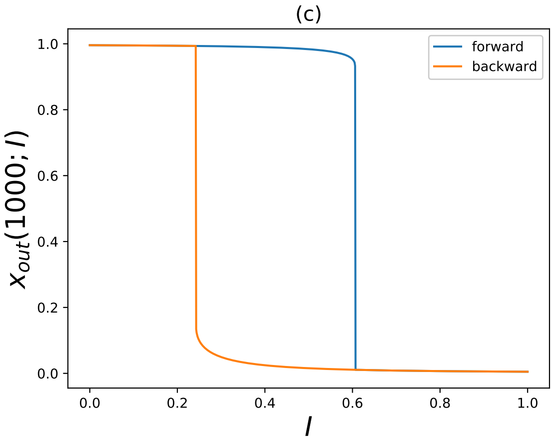


Sweeping I (no bifurcation case)



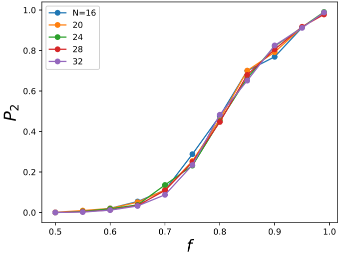
Sweeping I (saddle-node bifurcation case)

Responses of the fittest ensemble



Sweeping I (saddle-node bifurcation)

Appearance probability of two fixed points

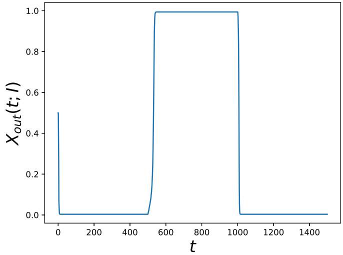


- Monotone increase against the fitness
  - Correspondence between the function and the number of the fixed points
- 99% of GRNs in the fittest ensemble have two fixed points

As the fitness increases, the big jump that the number of the fixed points changes takes place at somewhere in the course of evolution, irrespective of the evolutionary pathway

- Universality of evolution
- The fitness restricts the phenotype

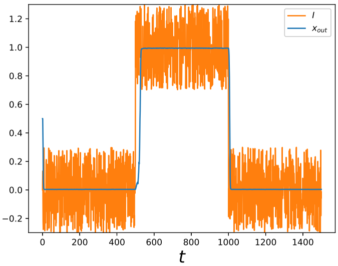
dynamical response



Dynamical response to sudden changes of the input

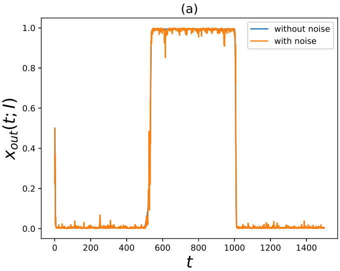
- 61% of the GRNs can respond properly.
    - Whether or not the bistable range include 0 or 1

Robustness against the input noise



- Number fluctuation of the input molecule
  - Uniform noise of $[-0.3, 0.3]$
- GRNs that can respond to the sudden change of input are robust against the input noise
  - The effect of the fixed-point switching

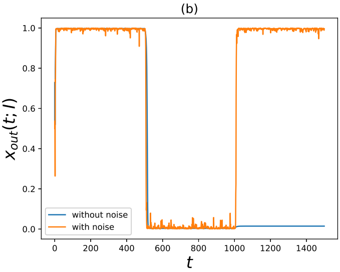
Robustness against the internal noise

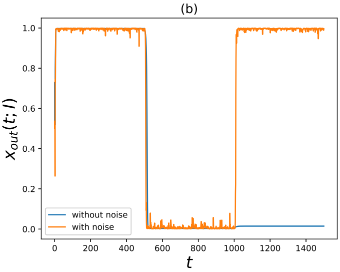


Dynamical response when the internal noise is applied

- Number fluctuation of TF
  - Uniform noises of $[-0.2, 0.2]$ are applied to all the input to each gene
- GRNs that can respond to the sudden change of input are robust against the internal noise

Noise-induced ultra sensitivity





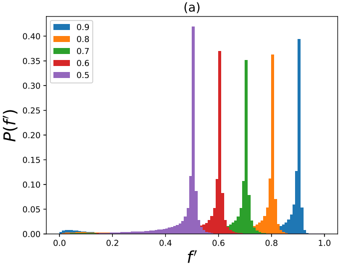
Fixed points and the robustness

- 60% of the GRNs having two fixed-points can respond properly to the sudden change of input
- They are robust against both input and internal noises
- Some of the GRNs exhibit the noise-induced ultrasensitivity
  - 70% in total can respond properly to the input

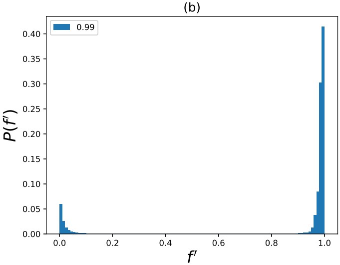
Mutational robustness

- Mutation of the single-edge deletion
  - A moderate mutation (e.g. slight change of TF)
  - We try all the possible mutations
  - Input/Output nodes are unchanged upon mutation

Distribution of the fitness f' after the mutation

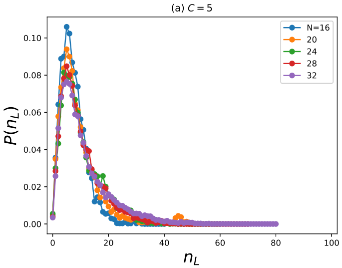


Several different f



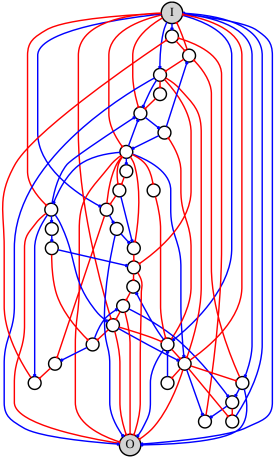
The fittest ensemble

- Majority of the edges are neutral against mutation
- For the fittest ensemble, most of the edges are either neutral or lethal
  - Intermediate edges are scarse



Distribution of lethal edges

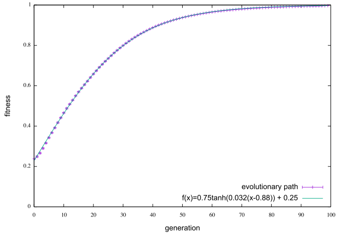
- Typical number of the lethal edges is 6
  - Independent of size
  - Larger GRNs are relatively robust
  - Some GRNs have no lethal edge



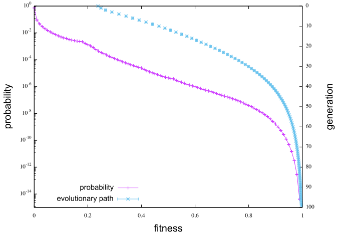
An example of GRN without a lethal edge

Results 2

- Compare the results of evolutionary simulations and the random sampling
- Slightly different model just for simplicity
  - Allow both self and mutual regulation. Fixed input/output nodes
- Population: 1000
  - Keep 500 samples from the highest fitness. Apply mutation to 500 copies
  - Perform 10000 runs and follow the evolutionary path of the fittest sample



Generation and fitness (fitting by tanh



Fitness landscape and the speed of evolution

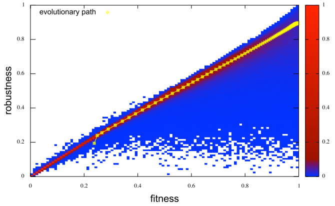
Speed of evolution is determined by entropy

Robustness index

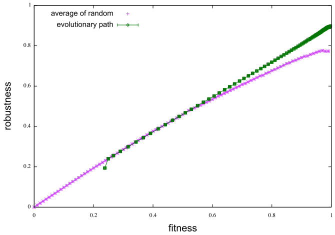
For all the possible deletion of edges,

$$\frac{1}{K}\sum_{edge} f'$$

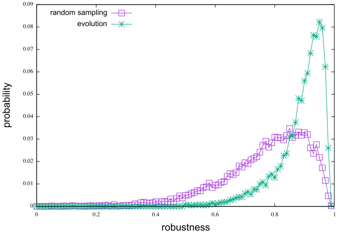
is defined as the robustness of each GRN



Robustness distribution and the evolutionary pathway



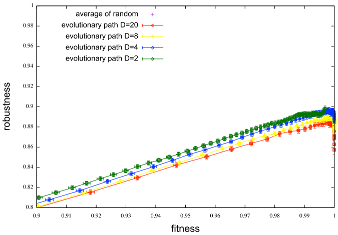
Avarage robustness and evolutionary pathway



- Distribution of the robustness for the fittest ensemble
- Distribution of the robustness for samples just after $f > 0.99$ is attained for the evolutionary simulations.

Evolutionary process is divided into two stages

1. Entropic stage
2. Robustness-aquiring stage



Pathways for different number of copies

Summary

GRNs of high fitness have the following features:

1. Ultrasensitivity (two stable fixed points)
    - A big jump irrespective of the evolutionary pathway
2. Three robustnesses
    - mutation
    - input noise
    - internal noise

Evolution enhances robustness