

遺伝子制御ダイナミクスの理解へ向けて

笹井理生

名古屋大学 工学研究科 計算理工学専攻

概要

細胞をコントロールする分子過程は大きな揺らぎにさらされており、それに伴った細胞状態の揺らぎを理解することは興味深い問題である。とくに、胚性幹細胞 (ES 細胞) の示す揺らぎ、iPS 細胞へのリプログラミングの揺らぎなど、細胞の揺らぎを理解することは応用上も大切な意味を持つ。こうした揺らぎを伴った細胞現象の背景にある遺伝子制御の物理、とくに非平衡キネティクスの原理を浮き彫りにしたい。非断熱性、渦流、非平衡ランドスケープと循環流など、非平衡過程の基本概念を探り、さらにゲノム運動の計算科学について紹介しながら、遺伝子制御ダイナミクスの理解への道筋を考える。

Toward Understanding Gene Regulation Dynamics

Masaki Sasai

Department of Computational Science and Engineering, Graduate School of Engineering, Nagoya University

Abstract

Cells are living in fluctuating molecular environment. Elucidation of these fluctuating mechanisms is important for biological applications as we can find examples in large fluctuations of embryonic stem (ES) cells or fluctuations of the reprogramming process to iPS cells. We here explain our efforts to elucidate the physical principles of gene regulation, which is a basic process to govern the fluctuating cell behaviors. By introducing basic concepts of non-equilibrium kinetics such as non-adiabaticity, eddy current, non-equilibrium landscape, and circulation current, and also by explaining computational methods to analyze genome movement, we discuss the perspective toward understanding gene regulation dynamics.

1 はじめに

遺伝子の産物である蛋白質や RNA の量を 1 細胞ごとに測る技術が普及するとともに、最近 10 数年の間に遺伝子の活動を調べる新しい分野が発展した。そこで判明したのは、遺伝子の活動は揺らぎの大きい確率現象であるということである。細胞が非平衡のメゾスコピック系であることを考えれば、こうした大きな揺らぎの存在は当然のことかもしれないが、なぜ大きな揺らぎにもかかわらず必要な生理過程は間違いなく進むのか、あるいは揺らぎを利用して進

むのか、その生物学的な意味は興味深い。こうした研究は細菌の細胞を対象として急速に進み、理論と実験が定量的に比較される統計物理の分野として成長した。しかし、ヒトやマウスなど高等生物を含む真核細胞は細菌とは大きく異なり¹、はるかに複雑な制御機構に基づいているため、真核細胞における遺伝子制御の理論はまだ殆ど手つかずの状態である。今回は、真核細胞における遺伝子制御ダイナミクスの理解に向けて、問題点と応用、今後

¹細菌は核をもたない原核細胞であり、ゲノム DNA は細胞内の広い範囲に分布している。真核細胞のゲノム DNA は細胞核の中に閉じ込められ、細胞質とは違った環境にある。

の展望を紹介してみたい。

2 遺伝子状態遷移の時間スケール

蛋白質の合成・分解により、蛋白質の量は確率的に変動していると考えられるが、もっとも単純な場合には、細胞内の蛋白質濃度 x の分布 $P(x, t)$ は、 t が大きい定常状態では、ガウス分布に接近するであろう。この分布を $P_{eq}(x)$ と書けば、 $U(x) = -\log P_{eq}(x)$ は放物線となるが、蛋白質濃度の確率的変動は、2次のポテンシャル面 $U(x)$ 上のブラウン運動とみなすことができる。

DNA 上の遺伝子制御部位に、アクチベーターと呼ばれるタイプの制御蛋白質が結合すれば遺伝子活動は活性化され、リプレッサーというタイプの制御蛋白質が結合すれば抑制される。すなわち、アクチベーターやリプレッサーの結合・解離により、遺伝子は活性状態と抑制状態の間を遷移すると考えられる。遺伝子状態の遷移速度と蛋白質濃度の変化速度の比を ω と書こう。 $\omega \ll 1$ 、すなわち、遺伝子状態の遷移速度が蛋白質濃度の変化速度よりずっと遅いときには、遺伝子が抑制されているときに限って考えた蛋白質濃度の定常分布を $P_-(x)$ 、活性化されているときの定常分布を $P_+(x)$ と書いて、蛋白質濃度はポテンシャル $U_- = -\log P_-(x)$ 、あるいは $U_+ = -\log P_+(x)$ のどちらかの上をブラウン運動していて、まれに U_- と U_+ の間を遷移するとみなすことができる (図1左)。一方、逆の極限、すなわち遺伝子状態の遷移速度が非常に速く $\omega \gg 1$ であるとき、 U_- と U_+ の間の遷移がブラウン運動より頻繁なため、蛋白質濃度は2つのポテンシャルが混ざった U_{\pm} を感じて、 U_{\pm} 上のブラウン運動とみなすことができる (図1右)。後者の $\omega \gg 1$ のときは、遺伝子遷移のダイナミクスを頭わに考えずに平均で置き換えたことになり、こうした扱いを固体物理や化学物理の用語を拝借すれば、断熱近似と呼ぶことができる。 ω は断熱パラメータであり、 $\omega < 1$ では非断熱的である。図1に示すように、 $\omega \gg 1$ の断熱極限と $\omega \ll 1$ の非断熱極限の扱いは比較的簡単であるが、その中間での非断熱領域 (図1中) は自明でない領域であり、 x の複雑な運動を渦流と呼んで経路積分表示を用いた分析がされている [1]。バクテリアでのアクチベーターやリプレッサーの結合・解離は一般的には秒から分の程度であり、蛋白質個数の変化に要する時間は細胞分裂時に大きく蛋白質の量が変化する時間、つまり数10分程度であると考えれば、 $\omega \gg 1$ であり断

熱近似が有効である。これが、バクテリアについて理論的な解析が容易であった理由のひとつである。

3 真核細胞の遺伝子スイッチング

真核細胞では、DNA はヒストンと呼ばれる蛋白質の複合体にまきつきヌクレオソームという単位をつくっており (図2)、ゲノムはヌクレオソームが数珠つなぎにつながったクロマチンの集まりとみなすことができる。細胞核のなかで、ヌクレオソームが比較的密に集まったヘテロクロマチンと呼ばれる領域へは、RNA 合成酵素やその他の因子が接近しにくく、その結果、遺伝子の活動は低下しているが、ヌクレオソームがゆるく集合しているユークロマチンと呼ばれる領域では遺伝子の活動は活発である。ヌクレオソームの集合の仕方のこうした違いは、ヒストンがどのように化学的に修飾されているかによって制御されている。

ヒストンの特定の場所がアセチル化されていると、アセチル基のマイナス電荷による静電反発のためにヌクレオソームは緩く集合し、遺伝子は活性化されるが、ヒストンの特定の場所がメチル化されていると、このメチル基を認識する蛋白質の働きでヌクレオソームは密に集合し、遺伝子は抑制される。こうしたヒストンの修飾状態は細胞分裂を越えて次世代の細胞に受け継がれるので、遺伝情報の次のレベルの情報という意味で、エピジェネティックな情報と呼ばれる。

真核細胞でヒストン状態の書き換えを行う酵素は、リプレッサーやアクチベーターなどの転写因子に引きつけられてヒストンに接近し反応するので、結局のところは転写因子が遺伝子を制御していると考え

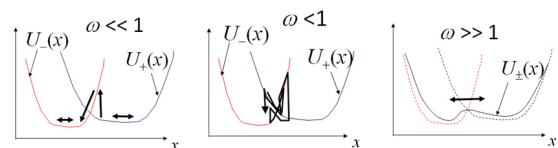


図1: 遺伝子スイッチの時間スケールにより異なるダイナミクスが現れる。蛋白質濃度を表す変数 x は、 $\omega \ll 1$ の場合、2つのポテンシャル面の片方をブラウン運動し、ときおり、もう片方のポテンシャル面に遷移する (左)。 $\omega \approx 1$ の場合、 x のブラウン運動と2つのポテンシャル間の遷移が混ざって、渦流をつくる (中)。 $\omega \gg 1$ の場合、 x は2つのポテンシャル面を平均した面の上をブラウン運動する (右)。

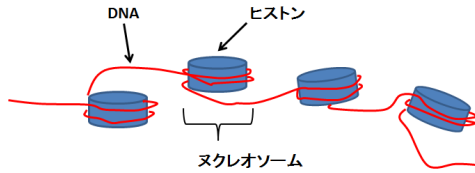


図 2: ヒストン蛋白質 8 量体のまわりに DNA が巻き付いて、ヌクレオソームと呼ばれる単位をつくる。ゲノムは、ヌクレオソームが数珠つなぎになったひもの集まりである。

れば、真核細胞でもバクテリアでも遺伝子制御のしくみは同じに思えるかもしれない。しかし、問題は時間スケールの違いである。ヒストン修飾変化の時間スケールは数日程度、転写因子の DNA への結合・解離は分程度の時間スケールであり、大きな開きがある。また、細胞の状態を直接規定する蛋白質の量の変化の時間スケールは数時間程度と考えることができる。ヒストン修飾変化の速度と蛋白質の量変化速度の比を ω_H と書き、転写因子の結合・解離速度と蛋白質の量変化速度の比を ω_T と書けば、 $\omega_H \ll 1 \ll \omega_T$ と表現できる。すなわち、ヒストン状態の変化は非断熱的であり、転写因子の結合状態変化は断熱的である。以下では、こうした時間スケールの階層性を考えることによって明らかになる現象について、いくつかのトピックスを紹介しよう。

3.1 iPS 細胞へのリプログラミングと非断熱ダイナミクス

図 3A のような遺伝子回路を考えてみよう。遺伝子 A の産物である蛋白質 A は遺伝子 A にとってはアクチベーターであるが、遺伝子 B の制御領域に結合してリプレッサーとして働く。遺伝子 B の産物である蛋白質 B は、自身にとってはアクチベーター、遺伝子 A にとってはリプレッサーである。この回路では A と B が拮抗しているため、蛋白質 A、B の分子数を N_A, N_B と書くと、 $N_A \gg N_B$ と $N_A \ll N_B$ の 2 状態が安定して存在すると想像できる。ここで、遺伝子 A を体細胞のあるタイプのもを特徴づける遺伝子、蛋白質 B を *Nanog* などの多能性細胞を特徴づける遺伝子と考えると、 $N_A \gg N_B$ は分化した細胞の状態、 $N_A \ll N_B$ を多能性細胞の状態とみなして、このモデルを分化やリプログラミングを表現するプロトタイプとすることができる。

2 つの安定状態に対応して、ランドスケープ

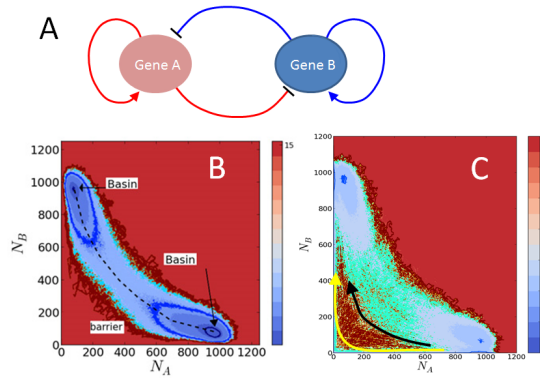


図 3: リプログラミングについての単純化された遺伝子回路。2 遺伝子からなるモデル (A)。この回路がバクテリアと同じような仕組みで働くとき、ランドスケープ $U(N_A, N_B)$ は $N_A \gg N_B$ と $N_A \ll N_B$ の 2 つの安定状態に相当したベイスンを持つ。リプログラミングは点線の峠道に沿った遷移に相当する (B)。ヒストンの遅いダイナミクスが転写因子の速いダイナミクスとカップルするとき、 $N_A \approx N_B \approx 0$ の新しいベイスンが現れ、ベイスン間をつなぐ新しい谷が出現する。この谷に沿ったリプログラミング (黄色い矢印) と山を回避する別の経路に沿ったリプログラミング (黒い矢印) が区別される。

$U(N_A, N_B) = -\log P_{eq}(N_A, N_B)$ には 2 つのベイスンが現れる。もし、この遺伝子回路がバクテリアと同じ仕組みで動くと考えたと、リプログラミングは図 3B の点線に沿った、2 つのベイスンをつなぐ峠道である。しかし、ヒストンの遅いダイナミクスを考慮すると、ランドスケープは大きい変更を受ける [2]。山中因子の働く機構の違いに応じて、ランドスケープ上の山を回避する 2 つの経路が可能で、それぞれキネティックな特徴が実験で区別できると予想される。

3.2 ES 細胞の揺らぎと非断熱ダイナミクス

同じ遺伝子を持つ ES 細胞のコロニーを同じ環境で培養しても、*NANOG* などの多能性維持にとって重要な蛋白質の個数は、細胞ごとに大きく異なる不均一性を示す。マウス ES 細胞で見られるこの大きな揺らぎは、2~3 日程度の時間で増減していると考えられ (細胞分裂は半日程度なので、それより長い時間で)、その機構は不明であった。バクテリアのように速い転写因子のダイナミクスを考えたモデルでも、ヒストンの遅いダイナミクスを考慮して $\omega_H \ll 1 \ll \omega_T$ と考えたモデルでも、説明がうまくできない。

4C法、Hi-C法などと呼ばれる最近の生化学的実験手法では、核内のクロマチンのどの領域とどの領域が接近しているかを網羅的に調査できる。こうした手法を用いると、ES細胞では *Nanog* などの多能性維持に必須な遺伝子のまわりにクロマチンが集合し、大きな構造をとっていることが示唆された。こうした大きなクロマチン構造の形成・崩壊は、細胞分裂の周期と同じ程度の時間を必要としているかもしれない。そこで、この速度と蛋白質量の変化速度の比を ω_A と書くことにすると、 $\omega_A < 1$ であることも想像できる。 $\omega_H \ll \omega_A < 1 \ll \omega_T$ と仮定したモデルは、実験で測定された NANOG の揺らぎを説明できる [3]。

3.3 エピジェネティックランドスケープと循環流

$\omega_H \ll \omega_A < 1 \ll \omega_T$ を仮定した遺伝子回路のモデルによって、多能性維持に必要な遺伝子群のヒストンが抑制状態に変化し、分化細胞に特有の遺伝子のヒストン状態が活性状態に変化する確率的な過程を追跡することができる。この過程で、それぞれの遺伝子から合成される蛋白質量によって細胞の状態を定義することができ、状態間の遷移確率を計算することができた。

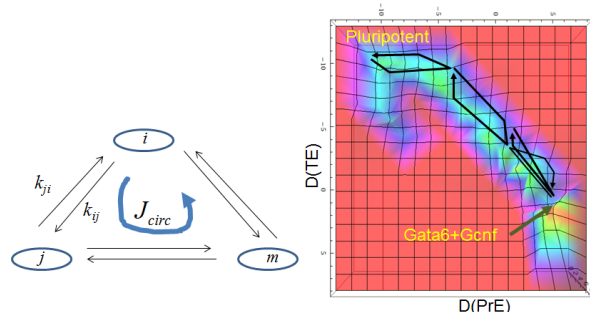


図4: 細胞分化のエピジェネティックランドスケープの計算。細胞状態 i, j, m 間のシミュレーション結果から、ランドスケープと循環流を計算することができる(左)。マウス ES 細胞からの分化のランドスケープ。横軸、縦軸はそれぞれある特定の細胞状態への接近度(右)。文献 [3] より転載。

この結果を理解するために、計算された確率過程から細胞状態変化のランドスケープ(エピジェネティックランドスケープ)を計算する。ただし、遺伝子回路のダイナミクスは非平衡過程であるため、詳細釣り合いは成り立っておらず、変化速度ベクトルはランドスケープの勾配に沿ってはいない。余分の循環

流を加えることで、整合性のある議論をすることができる(図4左)。この方法を用いて、ES細胞からの細胞分化の様子を表すエピジェネティックランドスケープと循環流を計算し、クロマチン運動の速度 ω_A の大小が分化に与える影響を解析した [3]。 ω_A が小さくて、ES細胞の揺らぎが大きいときに、ランドスケープは明確な構造を持ち、安定した分化を可能にすることが示唆された(図4右)。

4 展望

ES細胞のモデルの分析から、遺伝子回路のダイナミクスを理解するためには、クロマチンの構造ダイナミクスの理解が必要であることが示された。核内のクロマチンすべて、つまりゲノム DNA の運動を計算することは大きなチャレンジであるが、最近の Hi-C 法データを援用して可能になりつつある。図5は、ヒトやマウスよりずっと小さい、最も簡単なモデル真核細胞として多用される酵母の核内ゲノムの運動を計算した一コマである [4]。計算は、ゲノムが大きくランダムに動く様子を示して、ゲノム DNA の動きが遺伝子活動を制御する新しい機構の存在を示唆している。

ゲノム DNA 構造の運動を計算する計算科学、DNA 構造やヒストン状態を変数として取り入れた遺伝子ネットワークモデルの確率シミュレーション、非平衡確率過程を扱う統計力学、そうした理論の進歩が分子生物学、生物物理学の実験技術の進歩と協調して、細胞生物学の新しい概念と技術が産み出されると予想される。このように定量的に細胞を考える新しい分野には、思いもよらぬ発見が待ち受けているのではないだろうか。

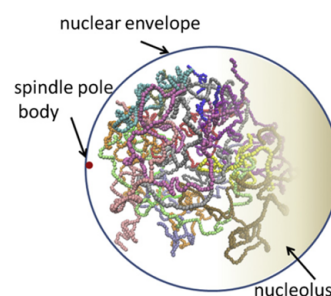


図5: 出芽酵母のゲノム動力学シミュレーション。16本の染色体が核内で液体中の高分子のように動く。文献 [4] より転載。

参考文献

- [1] K. Zhang, M. Sasai, J. Wang, Proc. Natl. Acad. Sci. USA **110** (2013) 14930.
- [2] S. S. Ashwin, M. Sasai, (2014)
<http://arxiv.org/abs/1410.2337>
- [3] M. Sasai, Y. Kawabata, K. Makishi, K. Itoh, T. P. Terada, PLoS Comput. Biol. **9** (2013) e100338.
- [4] N. Tokuda, T. P. Terada, M. Sasai, Biophys. J. **10** (2012) 296.