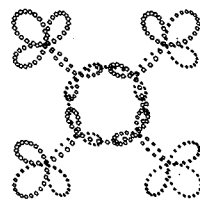


遺伝子発現の揺らぎの確率・ネットワーク解析

ロバストな遺伝子ネットワークの設計原理を目指して



富岡 亮太・小林 徹也・合原 一幸・木村 英紀

1. はじめに

生命を核酸 (DNA, RNA), 蛋白質とそれらの相互作用というミクロなレベルから説明するのが近年の分子生物学の立場である。これらの相互作用のなす複雑なネットワーク, ゲノム・プロテオーム系のダイナミクスの数理モデル¹⁻³⁾ は, さらに, 定量的に理解を深め, 動的な側面を明らかにすることに多くの成功を収めてきた。

分子生物学における実験技術の進歩が, これらの成功を後押ししている。近年, 多くの遺伝子の時空間的な発現パターンや, 遺伝子-蛋白質, 蛋白質-蛋白質間の相互作用ネットワークに関する網羅的な情報の取得・蓄積⁴⁾ が可能になってきた。また, DNA 操作技術の進歩により遺伝子ネットワークを設計し, それを実際に生体内に構築することが可能になっている⁵⁻⁷⁾。

分子生物学の世界でも, 生物をその構成要素に分解して詳しく性質を調べるという視点から, 多くの相互作用する機能単位のシステムレベルでの関係を理解する^{8,9)} という視点に転換しようという動きがある。また, 実験においてもモデリング, シミュレーション, 統合的なデータベースの利用によって, 生命現象の背景にある分子メカニズムを特定し, 効率的な計画に基づいて実験を行うことが現実的な方法論になってきた。このような「要素からシステムへ」の変化の一方で, 遺伝子ネットワークの振舞いを, 実際に設計し構築すること

で理解しようという動きが現れている^{8,10)}。これらは「分析から設計へ」の変化と見なすことができよう。生物が様々な外乱に対して恒常性を維持し, 正しく機能する背景には, どのような制御のメカニズムが必要か, あるいは, 所望の生理学的機能を実現するにはどのような設計が必要か, という生命の設計原理に関する問題に, 構成論的な方法で立ち向かうことが可能になりつつある。

2. 決定論から確率論へ

ゲノム・プロテオーム系のダイナミクスを記述する方法として, 化学反応速度論に基づく決定論的な記述が多く用いられてきた。さらに, 力学系としてのより深い数理的な理解も進んでいる。しかし, DNA 操作技術による構成論的なアプローチの進展は, 生物が必ずしも決定論的な振舞いをしないことを研究者に再認識させ, 「確率性」が再び多くの研究者の注目を集めている¹¹⁾。

Becskei¹²⁾, Ozbudak¹³⁾ らは遺伝子操作技術で遺伝子調節領域および蛍光蛋白質をコードする DNA 配列をプラスミド DNA 上に配置して大腸菌細胞内に導入し, 一つ一つの細胞の発光の強さを計測することで, 遺伝子の発現量の細胞集団内でのばらつきを調べ, 転写効率, 翻訳効率などのパラメータや, 負のフィードバックループなどのネットワーク構造を変えることで, 揺らぎをどう制御できるかを調べている。また, Ozbudak ら¹³⁾

は Langevin 方程式を用いた解析を行って、実験と数理モデルとの比較も行っている。

化学反応速度論はいわゆる熱力学的極限で化学成分の濃度の時間発展を記述する。他方で、有限でかつ極めて少数の分子 (DNA, mRNA, 調節蛋白質) の相互作用による時間発展は、決定論よりは確率論としての記述がより適当である。ゲノム・プロテオーム系の「確率性」の数理的な解析は、古くは、1970 年代に Berg¹⁴⁾ によって行われている。彼は、定常状態での揺らぎはポアソン分布より大きくなること、mRNA あたり翻訳される蛋白質の個数が幾何分布に従うことを明らかにした。また、細胞分裂による蛋白質の分配の揺らぎも考慮している。近年では、Arkin ら¹⁵⁾ は、大腸菌に感染するウイルス、 λ -ファージの発生経路の決定をマスター方程式を用いて定式化し、確率的シミュレーションを用いて実験結果をうまく説明した。そのほか、確率母関数^{16, 17)}、Langevin 方程式^{13, 18)}、Fokker-Planck 方程式^{19, 20)} などの手法が用いられ、急速に研究が進展しつつある。

遺伝子ネットワークの設計において、決定論的なモデリングのみで設計したネットワークが生体内で正しく動作することを保証することは難しい。単純に多くの蛋白質を生産することで、熱力学的な極限に近づけることも考えられるが、生体内資源の異常な消費は、生体の恒常性を破壊する危険性がある。これでは、たとえ興味深い現象を生体内に構築することができたとしても、実際に医療や実験に役立てることは難しい。一方で、実際の生命システムは、しばしば、非常に少数の分子で、ほとんど決定論的と言ってよい、信頼のできる振舞いを示す。また、生体が揺らぎを有効に活用しているという可能性もある²¹⁾。このような現象の背後にあるメカニズムを解明するには、「確率性」を本質的に考慮して数理モデルを作らなければならない。以下では、ゲノム・プロテオーム系における揺らぎを数理的に解析する筆者らの最近の取り組みを紹介する。また、確率的ネットワークグラフを用いた解析結果の情報表現、そこからのパターン発見の取り組みについても紹介する。

3. 揺らぎの数理解析

3.1 線形ノイズ近似

まずはじめに、遺伝子-蛋白質ネットワークを化学反応のダイナミクスとして記述する。システムは N 種類の化学成分、 M 個の化学反応 (R_1, \dots, R_M) からなるとする。各化学成分の個数を $\mathbf{X} = (X_1, X_2, \dots, X_N)^T$ で表す。化学反応速度論に従えば、 \mathbf{X} のダイナミクスは以下の常微分方程式で表される。

$$\frac{d\mathbf{X}}{dt} = A\mathbf{W}(\mathbf{X}). \quad (1)$$

ここで、行列 $A = \{A_i^j\}_{i=1, \dots, N, j=1, \dots, M} \in \mathbb{Z}^{N \times M}$ は *stoichiometric matrix* と呼ばれ、 A_i^j は、反応 R_j による i 番目の化学成分の個数の変化を表す。 $W_j(\mathbf{X})$ ($j = 1, \dots, M$) は *propensity function* と呼ばれ、 $W_j(\mathbf{X})\Delta t$ は、状態 \mathbf{X} にあるとき、微小時間 Δt の間に反応 R_j が起きる確率を表す。また、ベクトル $\mathbf{W}(\mathbf{X}) = (W_1(\mathbf{X}), \dots, W_M(\mathbf{X}))^T$ とする。定義に従えば A と $\mathbf{W}(\mathbf{X})$ の積は単位時間あたりの状態 (個数) 変化の期待値を表していることになる。

上の結果は、化学反応系の決定論的な記述である。システムに含まれる化学成分の個数が少ないとき、変化量の平均に従ってシステムが時間発展すると期待できなくなる。同じ初期値を持つシステムでも異なる時間発展をするであろうから、それらは確率過程からのサンプルパスと捉えられる。さらに、それらのなす分布 $P(\mathbf{x}; t) = \Pr(\mathbf{X} = \mathbf{x}; t)$ を考えれば、決定論的なダイナミクス (1) の代わりに確率分布の時間発展を表す方程式 (マスター方程式) が導出される²²⁾。

$$\frac{dP(\mathbf{x}; t)}{dt} = \sum_{j=1}^M \left[W_j(\mathbf{x} - \mathbf{a}^j) P(\mathbf{x} - \mathbf{a}^j; t) - W_j(\mathbf{x}) P(\mathbf{x}; t) \right]. \quad (2)$$

すなわち、ある状態 \mathbf{x} にある確率の変化率は、状態 \mathbf{x} にあって反応 R_j を起こして出ていく確率 (流出項) と状態 $\mathbf{x} - \mathbf{a}^j$ にあって反応 R_j を起こして入ってくる確率 (流入項) の和として表され

る。ここで、 \mathbf{a}^j は *stoichiometric matrix* A の第 j 列ベクトル、すなわち反応 R_j による状態変化を表す。

マスター方程式は、非常に一般的な記述を可能にしているが、一般に化学反応には2つ以上の化学成分が関わるため $\mathbf{W}(\mathbf{X})$ は非線形であり、さらに多変数であるために、マスター方程式を解析的に解く一般的方法は知られていない。また、数値的²³⁾にも多くのアルゴリズムが提案されているものの、計算量の点で多くのパラメータに対する依存性を明らかにすることは容易ではない。一方、定常状態では決定論的方程式(1)に従うが、その周りの揺らぎの解析が重要である場合も多い。線形ノイズ近似 (Linear Noise Approximation²²⁾) はそのような場合に強力な手法である。

決定論的方程式(1)の周りでの揺らぎが小さいとき、(1)の解を $\phi(t)$ とし(決定論項)、そのまわりの揺らぎを Ξ とし、 $\mathbf{X} = \phi(t) + \Xi$ と表すと、 $\mathbf{X} = \phi(t)$ の周りで線形化することにより^{*1)}、マスター方程式(2)は、 $\phi(t)$ の時間発展、

$$\frac{d\phi_i(t)}{dt} = A_i^j W_j(\phi(t)). \quad (3)$$

および、その周りの揺らぎ Ξ の確率分布 $\Pi(\xi; t) = \text{Pr}(\Xi = \xi; t)$ に関する線形 Fokker-Planck 方程式に帰着される^{*2)}。

$$\begin{aligned} \partial_t \Pi(\xi; t) = & \left[-\partial_{\xi_i} K_i^j(\phi(t)) \xi_j \right. \\ & \left. + \frac{1}{2} \partial_{\xi_i} \partial_{\xi_j} D_{ij}(\phi(t)) \right] \Pi(\xi; t). \end{aligned}$$

線形 Fokker-Planck 方程式の解は正規分布であり、その平均 $\mu = (\mu_1, \dots, \mu_N)$ と分散 $\Sigma = \{\Sigma_{ij}\}_{i,j=1,\dots,N}$ は以下のようなダイナミクスに従うことが知られている。

*1) 決定論項 $\phi(t)$ のオーダー Ω に対し、揺らぎ Ξ のオーダーは $\Omega^{1/2}$ であると仮定をする。

*2) $\partial_t = \partial/\partial t$, $\partial_{\xi_i} = \partial/\partial \xi_i$ とする。また、アインシュタインの記法を用いた。すなわち、上下に付いた同じ記号の添字はインデックスの全範囲にわたる積和を意味する。また、 ∂_{ξ_i} の添字は上付きと見なす。例えば、 $\partial_{\xi_i} K_i^j \xi_j = \sum_{i,j} \partial/\partial \xi_i K_i^j \xi_j$ 。

$$\frac{d\mu_i}{dt} = K_i^j(\phi(t)) \mu_j, \quad (4)$$

$$\begin{aligned} \frac{d\Sigma_{ij}}{dt} = & K_i^k(\phi(t)) \Sigma_{kj} \\ & + \Sigma_{ik} K_j^k(\phi(t)) + D_{ij}(\phi(t)). \end{aligned} \quad (5)$$

ここで、行列 $K(\phi) = \{K_i^k(\phi)\}_{i,k=1,\dots,N}$, $D(\phi) = \{D_{ij}(\phi)\}_{i,j=1,\dots,N}$ を以下のように定義した。

$$K_i^k(\phi) = \left. \frac{\partial A_i^j W_j}{\partial X_k} \right|_{\mathbf{X}=\phi},$$

$$D_{ij}(\phi) = A_i^k A_j^k W_k(\phi).$$

したがって、行列 $K(\phi(t))$ が $\phi(t)$ の時間発展に沿って安定であれば、 $\phi(t)$ の周りの平均および分散の時間発展(4),(5)は安定となる。特に、決定論項のダイナミクス(3)が安定平衡点 $\phi = \phi^{eq}$ を持つとき、その周りの揺らぎは、平均 0、分散を Lyapunov 方程式、

$$\begin{aligned} K_i^k(\phi^{eq}) \Sigma_{kj} + \Sigma_{ik} K_j^k(\phi^{eq}) + D_{ij}(\phi^{eq}) = 0 \\ (i, j = 1, \dots, N). \end{aligned} \quad (6)$$

の解とする定常分布に収束する。

まとめると、安定平衡点の周りの定常分布を求めるには、以下のような手順に従えば良い²⁴⁾。

1. 考えたいシステムの *stoichiometric matrix* A と *propensity function* $\mathbf{W}(\mathbf{X})$ を求める。
2. 決定論的方程式(3)の安定平衡点 $\phi = \phi^{eq}$ を求める。
3. 安定平衡点で行列 $K(\phi^{eq})$ および $D(\phi^{eq})$ を求める。
4. Lyapunov 方程式(6)を解く。

平衡点周りの揺らぎの評価は、例えば、遺伝子スイッチ、すなわち multi-stable でヒステリシスを持つネットワークを設計する²⁵⁾ 際にロバスト性の評価に用いることができる。また、この手法は多くの変数の揺らぎの間の相関を調べることができるため、負のフィードバック¹²⁾、二量体化^{26,27)}、背景分子との非特異的結合²⁷⁾ などの知られている揺らぎを抑制するメカニズムに対して、明快な解釈を与えることができる。

3.2 キュムラント発展方程式

マスター方程式のもう一つの見方として、分布の時間発展を直接求めるのではなく、分布をモーメント、キュムラント等の統計量で表現して、統計量の時間発展を求める方法がある。

マスター方程式(2)の両辺に X_i , あるいは $X_i X_j$ を掛けてすべての状態に関して和を取ると、それぞれ1次モーメント $\langle X_i \rangle$, 2次モーメント $\langle X_i X_j \rangle$ の時間発展を表す方程式を得る。

$$\frac{d\langle X_i \rangle}{dt} = A_i^j \langle W_j(\mathbf{X}) \rangle, \quad (7)$$

$$\begin{aligned} \frac{d\langle X_i X_j \rangle}{dt} = & A_i^k \langle W_k(\mathbf{X}) X_j \rangle + \langle X_i W_k(\mathbf{X}) \rangle A_j^k \\ & + A_i^k A_j^k \langle W_k(\mathbf{X}) \rangle. \end{aligned} \quad (8)$$

$W(\mathbf{X})$ が \mathbf{X} の2次以上の項を含む場合、1次のモーメントの時間発展(7)の右辺には2次以上のモーメントの項が含まれ、同様に、2次のモーメントの時間発展(8)の右辺には3次以上のモーメントが含まれる。一般に、 $W(\mathbf{X})$ が非線形な場合、 n 次モーメントの時間発展は $n+1$ 次以上のモーメントに依存することが示せる。したがって、ある次数までのモーメントに関して閉じた方程式を作ることはできない。当然、ある次数以上の依存性を打ち切るという近似を考えることができるが、一般には、ある次数以上のモーメントが小さくなると期待することはできない。そこで、モーメントではなく、キュムラントに関してある次数以上を打ち切ることを考える。キュムラントはモーメントに対して以下のように再帰的に定義される。

$$c_i = \langle X_i \rangle,$$

$$\begin{aligned} c_{i_1, i_2, \dots, i_n} = & \langle X_{i_1} X_{i_2} \dots X_{i_n} \rangle \\ & - \sum_{M \in \mathcal{V}'} \prod_{\{j_1, \dots, j_m\} \in M} c_{j_1, \dots, j_m}. \end{aligned}$$

ただし、 \mathcal{M} は集合 $\{i_1, \dots, i_n\}$ の一つの直和分割であって、 \mathcal{V}' は自明な分割 $\{\{i_1, \dots, i_n\}\}$ を除くすべての直和分割の集合を表す^{*3)}。ここで、

*3) 例えば、 $c_{i,j} = \langle X_i X_j \rangle - c_i c_j$, $c_{i,j,k} = \langle X_i X_j X_k \rangle - (c_i c_{j,k} + c_j c_{k,i} + c_k c_{i,j} + c_i c_j c_k)$ 。

n 個の添字を持つものを n 次のキュムラントと呼ぶ。

$W(\mathbf{X})$ は2次までの非線形性のみ含むと仮定し^{*4)}, $W_k(\mathbf{X}) = \alpha_k + \beta_k^l X_l + \gamma_k^{lm} X_l X_m$ ($k = 1, \dots, M$) と置く。仮定から、方程式(7), (8)に現れるキュムラントは、1次キュムラント(平均), 2次キュムラント(共分散)および3次キュムラントである。ここでは、3次のキュムラントが小さいと仮定して、1次, 2次のキュムラントの閉じた時間発展方程式を求めることを考える。これは、結果として確率密度関数の正規近似を与える。3次のキュムラントについて $c_{i,j,k} = 0$ ($i, j, k = 1, \dots, N$) と置くと、 $\langle X_i X_j X_k \rangle = c_i c_{j,k} + c_j c_{k,i} + c_k c_{i,j} + c_i c_j c_k$ と書くことができるから、(7), (8)は1次, 2次のキュムラントに関して閉じた方程式になる。

$$\frac{dc_i}{dt} = A_i^j \bar{W}_j(\mathbf{c}^1, \mathbf{c}^2), \quad (9)$$

$$\begin{aligned} \frac{dc_{i,j}}{dt} = & K_i^k(\mathbf{c}^1) c_{k,j} + c_{i,k} K_j^k(\mathbf{c}^1) \\ & + A_i^k A_j^k \bar{W}_k(\mathbf{c}^1, \mathbf{c}^2). \end{aligned} \quad (10)$$

ここで、 $\mathbf{c}^1 = (c_1, \dots, c_N)$, $\mathbf{c}^2 = \{c_{i,j}\}_{i,j=1,\dots,N}$ とする。行列 $K(\mathbf{c}^1) = \{K_i^k(\mathbf{c}^1)\}_{i,k=1,\dots,N}$ は1次キュムラントの関数、 $\bar{W}(\mathbf{c}^1, \mathbf{c}^2) = \{\bar{W}_k(\mathbf{c}^1, \mathbf{c}^2)\}_{k=1,\dots,M}$ は1次, 2次のキュムラントの関数であり、以下のように定義した。

$$K_i^k(\mathbf{c}^1) = A_i^j \{\beta_j^k + (\gamma_j^{lk} + \gamma_j^{kl}) c_l\},$$

$$\bar{W}_k(\mathbf{c}^1, \mathbf{c}^2) = \alpha_k + \beta_k^l c_l + \gamma_k^{lm} (c_{l,m} + c_l c_m).$$

ここで、関数 K は、前節での定義に等しく、 $K_i^k(\mathbf{c}^1) = \partial A_i^j W_j / \partial X_k |_{\mathbf{X}=\mathbf{c}^1}$ である。また、 $\bar{W}_k(\mathbf{c}^1, \mathbf{c}^2) = \langle W_k(\mathbf{X}) \rangle$ である。

キュムラント発展方程式²⁸⁾ (9), (10)を用いれば、平均値の時間発展だけでなく、分散の時間発展を(正規近似の範囲で)定量的に評価することができる。線形ノイズ近似と比べると、定常状態の近傍で $\mathbf{c}^1 = \phi$ と見なせるとき、 \bar{W} に2次キュムラント(共分散) $c_{l,m}$ が含まれるかどうかの違

*4) 3分子以上が同時に関わる反応は、連続する2つ以上の2分子反応で置き換えられるから、この仮定は一般性を損わない。

いとなる。特に、(9)は1次キュムラント(平均)の時間発展に対する2次キュムラントの寄与の存在を示している。これは $W(X)$ の非線形性に由来するもので、生体内の信号処理において感度の増幅等、積極的な役割を果たしているという指摘がある²⁹⁾。大数の法則に従えば、共分散項($c_{l,m}$)の寄与は平均に関して2次の項($c_l c_m$)に比べ一般に小さいが、これは分子の個数が数個のオーダーになると正しくないであろう。また、線形ノイズ近似は近似の良さが $\phi(t)$ の周りに分布がどれだけ集まっているか、すなわち(4),(5)が収束しているかどうか、に依存するため、安定な平衡点の周りでの定常揺らぎの評価では効果的であるが、非定常の確率的ダイナミクスを再現することは難しい。実際、大腸菌細胞は約60分の周期で成長および分裂を繰り返す。その際、蛋白質などの成分は2つの娘細胞に確率的に分配されると考えられている。キュムラント発展方程式は、細胞の成長による反応速度の変化を $W(X)$ の時間依存性として表現し、細胞分裂による成分の分配も適当な仮定の下に、正しく再現することができる。一方、キュムラント発展方程式は、非線形性のために一般に解析的に解くことは難しい。また、キュムラント間の依存性は複雑であるため、揺らぎを評価するだけでなく、揺らぎを生成、増幅、または、抑制するメカニズムを知るには、適切な情報表現の方法が必要となる。次の節では筆者らが提案する確率的ネットワークグラフを用いた揺らぎの伝播のグラフ表現について説明する。

4. 確率的ネットワークグラフ

確率的ネットワークグラフ²⁸⁾ G は以下のように定義される。 G のノード n_* は、 n_{c_i} , $n_{c_{i,i}}$, $n_{c_{i,j}}$ などのように、それぞれ、すべての化学成分に関する1次、2次のキュムラントを表す。有向エッジ $e_{(n_{c_a}, n_{c_b})}^k$ はキュムラント c_a のキュムラント c_b の時間発展への寄与を表し、エッジの重み k はその速度定数を表す。例えば、キュムラント $c_{i,i}$ に関するキュムラント発展方程式が、

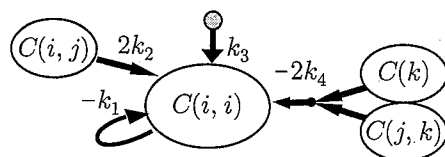


図1 キュムラント発展方程式(11)の確率的ネットワークグラフ表現。

$$\frac{dc_{i,i}}{dt} = -k_1 c_{i,i} + 2k_2 c_{i,j} + k_3 - 2k_4 c_{j,k} c_k \quad (11)$$

と表される場合、ノード $n_{c_{i,i}}$ へは、4つの有向エッジ $e_{(n_{c_{i,i}}, n_{c_{i,i}})}^{-k_1}$, $e_{(n_{c_{i,j}}, n_{c_{i,i}})}^{2k_2}$, $e_{(o, n_{c_{i,i}})}^{k_3}$, $e_{(\{n_{c_{j,k}}, n_{c_k}\}, n_{c_{i,i}})}^{-2k_4}$ が描かれる(図1)。ここで、 o は定数項を表すラベルなしのノードである。また、(11)の第4項は、ノード $n_{c_{j,k}}$ と n_{c_k} からのエッジを持つラベルなしノード(キュムラントの積を表す)および、そこから $n_{c_{i,i}}$ への重み $-2k_4$ のエッジで表現される。

確率的ネットワークグラフは、キュムラントの時間発展を追う際に強力なデータ表現手段となる。キュムラント発展方程式(9),(10)が求まれば、数値的に時間発展を求めることができる。その際、実際的には、時間発展に合わせてグラフの各エッジを寄与の正負で分類し、色分けして表示することができる。また、各エッジを寄与の相対的な大きさに応じて異なる太さで表示することもできる。適当に定めた閾値より相対的な寄与が小さいエッジは非表示にすることも考えられる。一般に成分間の相互作用が疎であることから、多くの共分散項($c_{i,j}$ ($i \neq j$))の大きさは無視できるほど小さいと考えられる。したがって、ほとんど寄与のない多くの経路が存在する。これらを相対的な寄与の大きさに評価することで、ネットワークを自動的に簡略化することができる。このようにして、確率的ネットワークグラフを用いてキュムラントの依存関係のなすネットワークを図示し、揺らぎがどのように伝播していくかを直感的に理解可能な形で表現することで、一種のComputer-Aided Analysisが実現される。これを図2に示す。このような解析法は、揺らぎを生成し、制御するメカ

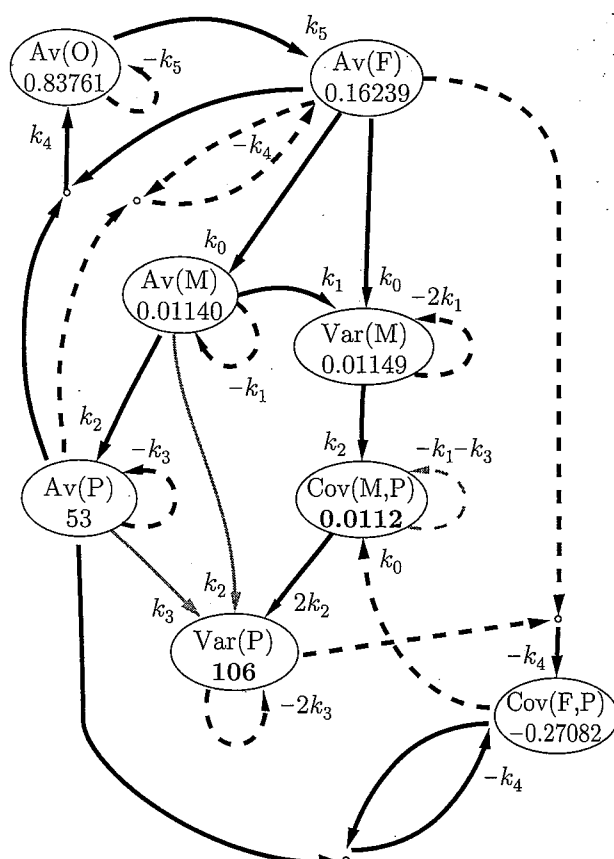


図2 転写調節による負のフィードバックを持つ遺伝子発現の確率的ネットワークグラフ。遺伝子発現に関わる部分だけを取り出して表示している。O, F, M, P はそれぞれ転写抑制された DNA, 転写可能な DNA, mRNA, 蛋白質を表す。ここでは視認性のため C の代わりに, Av, Var, Cov を用いる。これらはそれぞれ平均, 分散, 共分散を表す。ここでは実線が正, 点線が負の寄与を表し, 矢印の不透明度は相対的な寄与の大きさを表現する。Av(P) → Var(P), Var(M) → Cov(M,P) → Var(P), Var(P) → Cov(F,P) → Cov(M,P) → Var(P) の3つのパスがそれぞれ翻訳・蛋白質分解による揺らぎの生成, mRNA の揺らぎの伝播, そして蛋白質の揺らぎが負に伝播する効果を表す。

ニズムを解明する上で非常に大きなツールとなる。また、自動的なネットワーク簡略化と同時に揺らぎを生み出す特徴的なパターンを発見することも可能となる。ここで述べた方法は、系に含まれる化学反応を列挙し、行列 A , $W(X)$ で表現できれば、キュムラント発展方程式 (9), (10) の生成、時間発展の数値計算、確率的ネットワークグラフに

よるキュムラントのダイナミクスの表現、簡略化、パターン発見、という一連の解析をほぼ自動的に行うことが可能である。したがって、様々なモデルの振舞いを網羅的に調べ、実際の生理現象を良く説明するモデルを探したり、異なる設計原理を持つ遺伝子ネットワークの性能を比較したりする、といった解析を非常に手軽に実行することができる。これは、ゲノム・プロテオーム系の確率的ダイナミクスを扱う、新しいタイプのインフォマティクス研究の一つの基盤となり得るものである。

5. おわりに

本稿では、ゲノム・プロテオーム系と呼ばれる、遺伝子と蛋白質の複雑な相互作用からなるネットワークの確率的なダイナミクスを解析する筆者らの最近の取り組みを紹介した。まず、一般的な定式化を示し、次に、「線形ノイズ近似」、「キュムラント発展方程式」という2つの近似手法を紹介した。また、「確率的ネットワークグラフ」を用いた、解析結果の情報表現、自動的な簡略化による揺らぎの伝播の Computer-Aided Analysis について紹介した。ゲノム・プロテオーム系の研究分野では、急速な実験技術の進歩に対し、数理的な手法の発展が追いついていないのが現実である。以前から最もよく用いられて来た数理手法であるデータの統計解析については網羅的な測定技術の進歩とともに近年進展が目覚ましいが、まだ十分浸透しているとは言えない³⁰⁾。近年の構成論的な遺伝子ネットワークのダイナミクス研究の成功には力学系としての決定論的なモデリングが不可欠であったが、関係する分子の数の少なさに由来する確率性を考慮しなかったため、現実と解離してしまうこともあった。本稿で紹介した確率的なダイナミクスの解析・モデリングの手法は、確率・統計の手法と力学系の手法をうまく組み合わせた、新たな方法論である。「要素からシステムへ」そして「分析から設計へ」大きく舵を切ろうとしている生物学において、この方法論の有効性を具体的に検証していくことが今後の重要な課題である。

参考文献

- 1) 小林, 市瀬, 合原. ゲノム・プロテオームダイナミクスの数理. 数理科学, **39**, 48-54, 8月号 2001.
- 2) Tyson, J.J., Chen, K., and Novak, B. Network Dynamics and Cell Physiology. *Nat. Rev. Mol. Cell Bio.*, **2**, 908-916, 2001.
- 3) Hasty, J., McMillen, D., Isaacs, F., and Collins, J.J. Computational Studies of Gene Regulatory Networks: In Numero Molecular Biology. *Nat. Rev. Genet.*, **2** (4), 268-279, 2001.
- 4) Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumgarner, R., Goodlett, D.R., Aebersold, R., and Hood, L. Integrated Genomic and Proteomic Analyses of a Systematically Perturbed Metabolic Network. *Science*, **292** (5518), 929-934, 2001.
- 5) Gardner, T.S., Cantor, C.R., and Collins, J.J. Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, **403**, 339-342, 2000.
- 6) Elowitz, M.B. and Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature*, **403**, 335-338, 2000.
- 7) Atkinson, M.R., Savageau, M.A., Myers, J.T., and Ninfa, A.J. Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell*, **113** (5), 597-607, 2003.
- 8) Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. From molecular to modular cell biology. *Nature*, **402**, C47-C52, 1999.
- 9) Kitano, H. Systems Biology: A Brief Overview. *Science*, **295** (5560), 1662-1664, 2002.
- 10) Hasty, J., McMillen, D., and Collins, J.J. Engineered gene circuits. *Nature Insight*, **420**, 224-230, 2002.
- 11) Rao, C.V., Wolf, D.M., and Arkin, A.P. Control, exploitation and tolerance of intracellular noise. *Nature Insight*, **420**, 231-237, 2002.
- 12) Becskei, A. and Serrano, L. Engineering stability in gene networks by autoregulation. *Nature*, **405**, 590-593, 2000.
- 13) Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D., and van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nat. Genet.*, **31**, 69-73, 2002.
- 14) Berg, O.G. A Model for the Statistical Fluctuations of Protein Numbers in a Microbial Population. *J. Theor. Biol.*, **71**, 587-603, 1978.
- 15) Arkin, A., Ross, J., and McAdams, H.H. Stochastic Kinetic Analysis of Developmental Pathway Bifurcation in Phage λ -infected *Escherichia coli* Cells. *Genetics*, **149**, 1633-1648, 1998.
- 16) Thattai, M. and van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. USA*, **98** (15), 8614-8619, 2001.
- 17) Swain, P.S., Elowitz, M.B., and Siggia, E.D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA*, **99** (20), 12795-12800, 2002.
- 18) Simpson, M.L., Cox, C.D., and Sayler, G.S. Frequency domain analysis of noise in autoregulated gene circuits. *Proc. Natl. Acad. Sci. USA*, **100** (8), 4551-4556, 2003.
- 19) Hasty, J., Pradines, J., Dolnik, M., and Collins, J. Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. USA*, **97** (5), 2075-2080, 2000.
- 20) Kepler, T.B. and Elston, T.C. Stochasticity in Transcriptional Regulation: Origins, Consequences, and Mathematical Representations. *Biophys. J.*, **81** (6), 3116-3136, 2001.
- 21) Gammaitoni, L., Hanggi, P., Jung, P., and Marchesoni, F. Stochastic resonance. *Rev. Mod. Phys.*, **70** (1), 223-288, 1998.
- 22) van Kampen, N.G. *Stochastic processes in physics and chemistry*. Elsevier Science B. V., 1992.
- 23) Gillespie, D.T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.*, **81** (25), 2340-2361, 1977.
- 24) Tomioka, R., Kimura, H., Kobayashi, T., and Aihara, K. Multivariate Analyses of Noise in Genetic Regulatory Networks. preprint.
- 25) Kobayashi, T., Chen, L., and Aihara, K. Modeling Genetic Switches with Positive Feedback Loops. *J. Theor. Biol.*, **221** (3), 379-399, 2003.
- 26) Bundschuh, R., Hayot, F., and Jayaprakash, C. The Role of Dimerization in Noise Reduction of Simple Genetic Networks. *J. Theor. Biol.*, **220** (2), 261-269, 2003.
- 27) Morishita, Y. and Aihara, K. Noise-Reduction through Interaction in Gene Expression and Biochemical Reaction Processes. *J. Theor. Biol.* submitted.
- 28) Kobayashi, T. and Aihara, K. How Does Noise Propagate in Genetic Networks? A New Approach to Understand Stochasticity in Genetic Networks. preprint.
- 29) Paulsson, J., Berg, O.G., and Ehrenberg, M. Stochastic focusing: Fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl. Acad. Sci. USA*, **97** (13), 7148-7153, 2000.
- 30) Tilstone, C. DNA microarrays: Vital statistics. *Nature*, **424**, 610-612, 2003.

(とみおか・りょうた, 東京大学大学院新領域創成科学研究科)
 (こばやし・てつや, 東京大学大学院新領域創成科学研究科)
 (あいはら・かずゆき, 東京大学生産技術研究所)
 (きむら・ひでのり, 東京大学大学院新領域創成科学研究科)