

# 数理科学は生物学に革新をもたらし得るか？

理化学研究所

望月 敦史

## 1. 生物学から数理科学に向けられた期待

近年の生物学のめざましい発展により、さまざまな生命現象のメカニズムがつぎつぎと明らかにされてきた。分子遺伝学や生化学といった方法により分子レベルの知見が蓄積した結果、生物の機能には様々な遺伝子や蛋白質などの生体分子がかかわっていることや、それらの制御の関係が複雑なネットワークを形作っていることが、分かってきた。飛ぶ鳥を落とす勢いの生物学に見えるが、実はここ 10 年ほど新たな展開を求めて手法を模索し続けている。これまでの生物学は、実験事実を少しずつ積み重ねて行く実証研究に支えられてきた。その結果、生体分子の複雑なシステムが明らかになったが、一方でそれらシステムから、どのようにして生命機能が作り出されるのか、まだ十分分かっていない。細胞がもつ自律性や恒常性、発生の過程で作り出される機能的な形、環境変動に対し行動を変える適応性、これらの「いかにも生命らしい働き」が生まれるメカニズムは、今も変わらず生物学の大きな謎である。

生命現象の本質は、生体分子システムが作り出す活性のダイナミクスや空間パターンであると考えられている。つまりネットワークに基づき、生体分子の活性ダイナミクスが走り、それらが生み出す定常状態の多様性や、周期振動などの動的挙動こそが、生物らしい振る舞いの起源なのだと考えられている。これに対し、従来の個別的で還元論的視点ではなく、複雑なシステム全体を統合的に捉え、そのダイナミクスを理解する新たな方法が、実験生物学者の側から求められている。特に数理科学や物理学などの理論的な手法は、大きな期待を集めている。

一方で、実験生物学者ばかりでなく、数理科学者あるいは物理学者のあいだでも、ここ数年、変化がみられるようになってきた。単に従来の数理理論を適用する対象として生命現象を捉えるのではなく、生物学の問題そのものの解明に向けて理論的に取り組む研究が、あらわれてきた。実験生物学者と数理科学者との共同研究も積極的に進められるようになってきている。理論研究者の数もしだいに増し、人的なネットワークや新たな研究領域を形成しつつある。

現在、生物学から理論科学に特に期待されていることは、「予測を与える理論」である。既知の事実から数理的に未知の予測が導かれ、実験的検証によって事実が明らかにされる、

その繰り返しによって進むような新しい生物学の到来が、実験生物学者から望まれている。実のところ、観察された現象を後追いで説明する理論であれば数多くあるが、実験に先立って予測を与えるような理論は未だ限られている。また、そのような理論を作り出すためには、古典的な物理学とは少し異なる考え方が必要かもしれない。というのは、機能の実体であるタンパク質について、その物理的性質が良く分かっていない段階で、それでも遺伝子の働きが実証的・現象論的に示されている、という状況が現在の生物学だからである。私は、生物学で有効に機能する理論は、むしろ論理学に近いものではないかと考えている。

## 2. 生命の複雑制御ネットワーク

以下では例として、生物学の基本となる遺伝子の制御について説明しよう。我々の体を構成する細胞の一つ一つが、デオキシリボ核酸（DNA）の高分子を持っている。遺伝子とはその DNA の上に存在する機能単位である。遺伝子には活性状態と不活性状態があると考えられており、活性化した遺伝子からは（mRNA という中間産物を経て）、タンパク質が作られる。DNA の塩基配列という形で遺伝子が持っている情報は、そのままタンパク質のアミノ酸配列に翻訳される。アミノ酸配列はタンパク質の立体構造、さらにはその機能を決定していると考えられ、これらの性質は元をたどれば DNA の配列にさかのぼることができる。生成されたタンパク質の一部は、転写因子として機能する。つまり遺伝子配列の近傍に結合することで、遺伝子の活性／不活性を切り替える。活性化／不活性化を切り替える転写因子の種類や結合の規則は、遺伝子ごとに決まっていると考えられている。遺伝子の活性の結果作り出されるタンパク質が、再び遺伝子の活性／不活性を制御するという意味で、これは閉じた力学系として考えることができる。

多くの生物現象に多数種の遺伝子が関わっており、それらは複雑な制御関係を持っている。実験的に明らかにされた多くの制御を理解するために、それらをまとめたネットワークが用いられてきた。例えば Fig. 1 はホヤの初期発生の 8 細胞期から尾芽胚期にかけて、13 種類の細胞の違い（細胞分化）が作り出される際に働く 80 個あまりの遺伝子と、それらの間の制御関係を示しているグラフである<sup>1)</sup>。尾芽胚はおたまじゃくしのような形をしており、それまでの非常に速い発生が一段落する時期でもある。細胞の分化状態は、活性化した遺伝子の違いだと考えられる。基本的に全ての細胞は同じ遺伝子のセットを持っており、同一の遺伝子制御ネットワークから、異なる複数の遺伝子活性状態が作られる。細胞運命が分かれていく過程は、細胞質因子や細胞間相互作用の違いがトリガーとなり、遺伝子活性のダイナミクスが互いに異なる状態に進んだのだ、と理解される。つまり Fig. 1 の制御ネッ

トワークは、それに基づく遺伝子活性のダイナミクスが、多様な定常状態を作り出す性質を備えているのだろう、と期待できる。

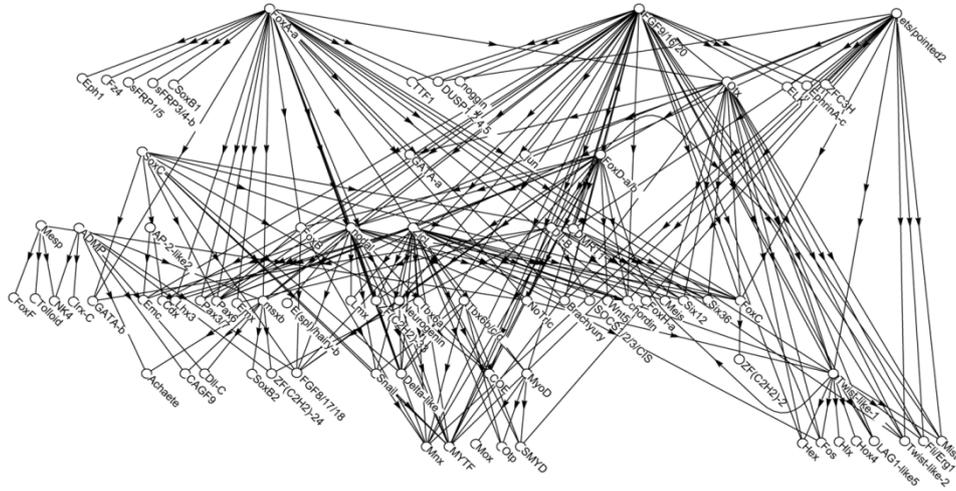


Figure 1 ホヤの発生で働く遺伝子制御ネットワーク (文献1より改変)

この制御のグラフは、主に攪乱実験によって、一つ一つ導かれたものである。例えば特定の遺伝子 A の機能を実験的に破壊し、その状態で他の遺伝子への影響を観測することができる。遺伝子 A の破壊により、本来活性であったはずの遺伝子 B が不活性化したならば、もともと遺伝子 A から遺伝子 B への活性化の制御があったのだ、と推測できる。

ネットワークの構造とそこから生まれるダイナミクスとの関係を明らかにすることが、生物学における最重要課題の一つである。しかし、この様な方法で決められた制御ネットワークは、明らかにダイナミクスを決めるための十分な情報を持っていない。制御ネットワークが与える情報は、システムに含まれる変数の間の依存関係だけである。構造とダイナミクスの関係が理解されていない一方で、遺伝子間の相互作用の知見はますます進み、ネットワークは複雑化する一方である。

### 3. ネットワークの構造とダイナミクス

以下は遺伝子のネットワークシステムに対して、私が進めた数理的研究を簡単に紹介しよう。遺伝子ネットワークの動態は、遺伝子の活性  $z_k$  の変化を記述する以下の微分方程式系で一般的に理解される<sup>2,3)</sup>

$$\dot{z}_k = F_k(z_k, z_{l_k}), \quad k = 1, \dots, N. \quad (1)$$

ここで  $k$  は遺伝子を区別し、 $F_k$  は分子活性の時間変化を記述する関数である。ただし上で述べたように、多くの生物システムで  $F_k$  の形は未知である。ここで(1)の動態は、制御ネ

ネットワークに対応する有向グラフ上のシステムだと考える。グラフの頂点  $i$  から頂点  $k$  への辺 ( $i \rightarrow k$ ) は,  $i \in I_k$  となる  $(i, k)$  のペアを与える。言い換えるならば,  $F_k$  の引数である  $I_k \subseteq \{1, \dots, N\}$  は, 遺伝子  $k$  を制御する遺伝子の集合に相当する。ここで  $F_k$  に対する制約として, 一つだけ仮定 (分解条件) をおく, すなわち

$$\partial_1 F_k(z_k, z_{I_k}) < 0. \quad (2)$$

ここで  $\partial_1$  は  $F_k$  の最初の引数  $z_k$  に関する偏微分を意味する。生体分子の活性の実体が mRNA やタンパク質などの物質である以上, 減衰や分解は必ず働く。もし  $k \notin I_k$  であれば, つまりグラフが自己制御ループを含んでいなければ, 分解条件から  $z_k$  自身の効果は抑制的である。もちろん,  $k \in I_k$  であれば, つまりグラフが自己制御ループを含んでいれば, そうとは限らない。ただし, 自己制御が働いていても抑制的にしか働いていないようであれば, 自己制御ループがないことと同様である。

制御ネットワークが与える情報は, 系に含まれる変数の間の依存関係, 言い換えるならば方程式系の引数の情報だけである。であれば, 引数の情報だけからダイナミクスについてどれだけの情報が引き出せるか, という問題を考えれば良いのではないか。このアイデアに基づき我々は, ネットワークだけから情報を引き出す理論を導き出した<sup>2,3)</sup>。これは, (1)で記述される微分方程式系に対して, 一般的に成り立つ。以下ではその主要な結果だけ説明しよう。この理論は数学の異なる分野の二つの概念, グラフ理論における Feedback vertex set と, 力学系における Determining nodes とを結び付けるものである。Feedback vertex set とはグラフの頂点の部分集合であり, 「それらを除くとグラフからループが無くなるようなもの」のことをいう<sup>4)</sup>。もう一つの Determining nodes は, 以下のように定義される変数の部分集合である。任意の二つの解軌道  $z(t)$ ,  $\tilde{z}(t)$  について,

$$\tilde{z}_I(t) - z_I(t) \xrightarrow{t \rightarrow +\infty} 0 \quad (3)$$

ならば,

$$\tilde{z}(t) - z(t) \xrightarrow{t \rightarrow +\infty} 0 \quad (4)$$

であるとき,  $I \subseteq \{1, \dots, N\}$  を Determining nodes とよぶ。

これはもともと流体力学の文脈で議論されていた概念で, (系の全ての変数を観測しなく

とも) それらを観測すれば系全体の振る舞いを観測したことに相当するような変数の部分集合である<sup>5)</sup>. 本来 Navier-Stokes 方程式の解を決定するための, 計測すべき空間上の有限の点に関する概念であり, グラフとの関わりは全く議論されてこなかった. 今回我々が証明したことは, 「ネットワーク上のダイナミクスについては, ネットワークの Feedback vertex setこそが determining nodes だ」ということである<sup>2)</sup>. 証明は文献 2 に示すが, 議論の本質は関数の一意性を駆使することにある.

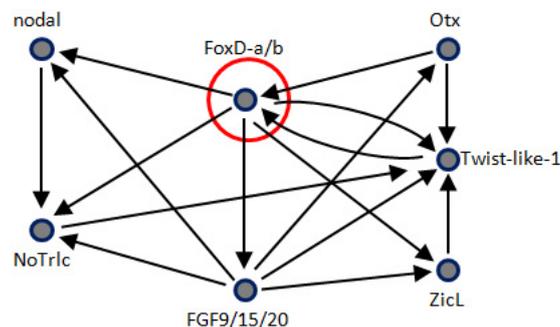


Figure 2 ホヤのネットワークの解析. Feedback vertex set を○で囲った.

ここでは, 生物のネットワークを使って, 具体的にこの理論の有用性を説明しよう. ホヤのネットワークは一見複雑そうであるが, 「何からも制御されない遺伝子」と「何も制御しない遺伝子」を順次除くと, 7つの遺伝子からなるネットワークの中心部分が得られる (Fig. 2). 注意深く見ると, このネットワークの Feedback vertex set はわずか 1 つの遺伝子 FoxD-a/b だけを含む. このネットワークに基づき作りだされる力学的挙動は, 必ずこの遺伝子の挙動として観察されるはずである. つまりネットワークが正しければ, 13 通りの細胞分化状態は, FoxD-a/b 遺伝子の活性の違いとして捉えられるはずである. ここで注意してほしいのは, 発生生物学の文脈では, 遺伝子活性状態はしばしば活性または不活性の二値で捉えられることである. もしこの理解が正しければ, ホヤの遺伝子ネットワークは, たった 2 状態しか作り出すことができない. この結果を生物学的に解釈する方法は 2 通り考えられる. つまり, 遺伝子活性を二値で捉えることに無理があるか, 或いはネットワークが不完全であるかのいずれかである. 前者を考えた場合, FoxD-a/b の活性をより詳細に連続量として定量する必要がある. 詳細に調べた FoxD-a/b の活性は 13 通りの細胞分化状態を反映しているかもしれない. 或いは後者に示すように, 現在のネットワークでは細胞分化の多様性を説明するには不十分であり, 未発見の制御が残っているのかもしれない.

このように我々は, ネットワークに基づいて作りだされる力学的挙動は一部の変数のみで捉えられること, またそれらの変数はネットワークの構造だけから決定できることを示

した。ネットワークの構造によっては、観察される生物学的振る舞いを実現できる能力が無いことが分かり、未知の制御の存在を予測することもできる。

#### 4. おわりに

最後に、生物学の理論が今後向かうであろう方向について、簡単に述べたい。生物現象を数理的に解明しようとする際に、仮定を導入し数理モデルを構築するのが、現在の通常の方法である。したがって、モデルの解析により、特別な結果が得られたとしても、モデル化の際に導入された仮定を、まず疑う必要がある。理論生物学者は常にその点に注意を払って研究しているが、それは、理論・数理の立場から実験生物学に対し、強く予測を提示することが難しい理由にもなっている。今回紹介した理論では、制御関数の形やパラメータの値などの仮定を導入する必要が全くない。したがってこの論法で得られた結果は、直接実験へ還元できる。例えば何か矛盾が見つければ、そこから実験情報を疑うことが可能であり、理論の側から極めて強い提言が可能である。このような理論は何も我々の研究だけに限らない。また、実のところこの理論は、分子生物学の世界で無意識かつ直感的に用いられてきた論法を元に、数学的に定式化することから生まれたものである。定式化することで強力な理論となるようなアイデアは、生物学にはまだ多数隠れているように思われる。生物学はそこで見られる現象の豊かさだけでなくその論法においても、これからの理論科学に多くの材料を与えてくれるだろう、と期待している。

#### 引用文献

- 1) Imai, K. S., Levine, M., Satoh, N. and Satou Y. (2006). Regulatory Blueprint for a Chordate Embryo. *Science* **312**, 1183-1187.
- 2) Fiedler B., Mochizuki A., Kurosawa G. and Saito D. (2013) Dynamics and control at feedback vertex sets. I: Informative and determining nodes in regulatory networks. *J. Dyn. Differ. Eqns.* 25, 563-604.
- 3) Mochizuki A., Fiedler B., Kurosawa G. and Saito D. (2013) Dynamics and control at feedback vertex sets. II: A faithful monitor to determine the diversity of molecular activities in regulatory networks. *J. Theor. Biol.* **335**, 130-146.
- 4) Akutsu, T., Kuhara, S., Maruyama, O. and Miyano, S. (1998). A system for identifying genetic networks from gene expression patterns produced by gene disruptions and overexpressions. *Genome Informatics* **9**, 151-160.
- 5) Foias, C. and Temam, R. (1984) Determination of the solutions of the Navier-Stokes equations by a set of nodal values. *Math. Comput.* **43**, 117-133.