

増殖過程の統計物理的構造

小林 徹也

概要

一般の物理系と異なり、細胞や個体などの生物集団は、自己の状態を変異させ、また自身の複製を生成することで増殖をすることができる。この変異と増殖のダイナミクスは生命進化を司る基本過程であり、その理解は非生物系と生物系の共通構造および本質的な差異を解明するためにも必須である。また、生物系は積極的に環境の情報を内部に取り込みそれを処理することにより、集団としての適応度(増殖率)を制御することができる。適応度と情報の関係を理解することは、我々の脳のような高度な情報処理機構が進化の過程でどのように選択されてきたのかを明らかにするためにも重要である。

この問題に関し、本発表では増殖過程の有する数理構造に着目する。具体的には、増殖ダイナミクスの経路積分表現とそれに伴う遡及的表現を導入することにより、増殖過程に内在する統計物理的構造を明らかにする。この構造を用いることにより、確率熱力学と同様に、増殖集団の適応度などのマクロな諸量に成り立つゆらぎ関係を示す。また、この表現を活用することで、適応度と情報に成り立つ交換関係を、統一的に明らかにする。これらの生物学的な意義を議論するとともに、進化の問題への他の物理的アプローチの可能性についても言及をする。

1 導入

1.1 表現形選択と適応

細胞を始めとした生体システムは複雑で時間的に変動する環境下で生存している。しばしば環境の変動は確率的で、また細胞などに対して致命的な環境が生じることもある。このような状況に対し、細胞やより高次の個体などは複数の表現形を使い分けることによって対応していることが知られている。表現形とは細胞や個体の形質を指す概念であり、例えばタンパク質の発現量、細胞のサイズ、などに対応する。遺伝子が同じでもこのような表現形は細胞ごとに確率的に異なる。細胞はこのような確率性から生じる表現形の多様性を生存戦略としても活用している [1]。

例えばよく増殖する表現形とあまり増殖しない表現形の2つを持つ細胞を考える。この細胞の集団が増殖に適した環境にいる場合、よく増殖する細胞が集団のサイズを指数的に拡大する。一方、この細胞の集団に抗生物質などの致命的な薬剤が加えられた場合を考える。抗生物質は「細胞を殺す」と言われるが、多くの場合、その作用機序は細胞が増殖にとまって外界の物質を取り込む際に細胞に取り込まれ、代謝を始めとした細胞の生存や増殖に重要な経路に作用して細胞集団の増大を抑制する [2]。したがって、増殖しない細胞に対しては、あまり抗生物質は作用しない¹。さらに一部の薬剤は、細胞内の特定の酵素などとの相互作用で機能を発揮する場合があります、この場

¹完全に作用機序によるので常にそういうわけではない。

合はその酵素の確率的な遺伝子発現によって細胞ごとに薬剤の効果が大きく異なりうる [3]。遺伝的変異を獲得することなく細胞が薬剤耐性を生き延びる現象は persistence 現象として古くから知られており [4]、このようなあまり薬剤が効かない細胞が存在することにより、特に薬剤耐性菌の問題や [5]、癌の薬剤治療などで問題となっている [6, 7]。

このような増殖を介して、生物個体集団と環境とが相互作用する過程は、一般的に自然選択もしくは選択と呼ばれる。また確率的な表現形の選択とそれによる生存確率 (適応度) の集団としての上昇は、細胞の両賭戦略 (bet-hedging) と呼ばれる [8, 9, 10, 11, 12, 13, 14]。

細胞集団が環境により適応した戦略を採用するという事は、何らかの意味で環境に適合した戦略を選択することである。そのため、環境に適合した戦略は環境のある種の情報を反映することが予想される。また、そのような最適な戦略で達成しうる増殖率 (適応度) は、環境変動の統計特性と細胞自体の増殖特性によって定まることから何らかの相互関係が存在することが期待される。このような関係性は、細胞増殖、さらに広くは進化に内在する構造・制約として古くから注目されてきているものである。

1.2 情報と適応度

さらに細胞が環境状態に関する部分情報を、レセプターなどの環境感知系を用いて取得できる場合、表現形の出現確率を情報に合わせて適切にバイアスすることにより、より適応度を増加させることが可能になる [15]。この適応度の上昇は、環境の情報を得ることによる直接的な対価であり、生体システムにおける情報の価値を考える上での重要なフレームワークである [16, 17, 18, 19, 20]。またより長時間的に見た場合には進化などの問題にも関連する [21]。得られた情報によりどれだけの適応度上昇が見込まれ、その限界は何なのか、そして環境の感知して得られる情報は、自然選択を介した環境と細胞の相互作用によって得られる情報とどのように関係するのか? は進化論における極めて重要な問題であるといえる。

以下では、細胞の表現形選択、増殖、そして環境感知に関わるモデルを導入し、増殖系の持つ構造と、細胞にとっての情報の価値を統計物理的な視点から明らかにする。

2 細胞増殖のモデル

細胞の集団は種類の表現型を持ち、細胞の表現型を $x \in \mathcal{S}_x$ で表す。時刻 τ における表現形 x を持つ細胞の集団数を $\mathcal{N}_\tau(x) \in \mathbb{R}_+$ とする。細胞数は本来は正の整数であるが、実数で近似できる程度に集団サイズが大きい場合を考える。簡単のため、離散的な時刻 $t \in \{0, 1, \dots, \tau\}$ を考える²。時刻 $t = 0$ における細胞数が \mathcal{N}_0 であり、かつ時刻 t における表現形 x の細胞の増殖率が $e^{k_t(x)}$ である時、各細胞がその表現型を変えないとすると、

$$\mathcal{N}_t(x) = e^{k_t(x)} \mathcal{N}_{t-1}(x) = \left[\prod_{s=0}^{t-1} e^{k_{s+1}(x)} \right] \mathcal{N}_0(x)$$

となる。persistence を始めとする現象では、細胞は確率的に表現型を変える。それを表現するため、時刻 t に表現型 x であった細胞が時刻 $t + 1$ に表現型 x' に表現型を変化させる確率を

$$\mathbb{T}_{t+1}(x'|x)$$

²技術的な複雑化は伴うが、状態の連続化、時間の連続化も通常の方法で可能である

とする。すると、時刻 t において $\mathcal{N}_t(x)$ の細胞集団からは、 $\mathbb{T}_{t+1}(x'|x)\mathcal{N}_t(x)$ の表現形 x' 持つ細胞集団が生まれる。よって、

$$\mathcal{N}_{t+1}(x') = e^{k_{t+1}(x')} \sum_x \mathbb{T}_{t+1}(x'|x)\mathcal{N}_t(x) = \sum_x \mathcal{A}_{t+1}(x'|x)\mathcal{N}_t(x)$$

が得られる³。ここで $\mathcal{A}_{t+1}(x'|x) := e^{k_{t+1}(x')}\mathbb{T}_{t+1}(x'|x)$ である。

2.1 環境の影響

さらに細胞が存在する環境は時刻ごとにその状態が変動する。 $(t-1, t]$ の間の環境状態を $y(t) = y_t \in \mathcal{S}_y$ と表す。環境の状態も離散であるとする。 $k_t(x)$ の時間依存性は主にこの環境状態に依存し、 $k_t(x) = k(x, y(t))$ と表される。同様に一般に細胞も環境の状態を何らかの方法でモニターしてその表現型を変えられる。細胞が時刻 t において環境を観測して得た情報を z_t とする。 z_t は一般に y_t と相関し、したがって環境についての情報を持つ。 $\mathbb{T}_t(x'|x) = \mathbb{T}_t(x'|x; z_t)$ となる。この場合

$$\begin{aligned} \mathcal{N}_{t+1}(x') &= e^{k(x', y_{t+1})} \sum_x \mathbb{T}(x'|x, z_{t+1})\mathcal{N}_t(x) \\ &= \sum_x \mathcal{A}(x'|x; y_{t+1}, z_{t+1})\mathcal{N}_t(x), \end{aligned}$$

を得る。

2.2 適応度と経路積分表現

ここで細胞集団の適応度 Ψ を

$$\Psi[\mathcal{Y}_\tau] := \ln \frac{\sum_{x'} \mathcal{N}_\tau(x')}{\sum_{x'} \mathcal{N}_0(x')}$$

と定義する。適応度は、初期時刻 $t=0$ から $t=\tau$ までに細胞集団が増大した指数増大率に対応する⁴。今、経路 $\mathcal{X}_\tau := \{x_t | t \in [0, \dots, \tau]\}$ および $\mathcal{Y}_\tau := \{y_t | t \in [0, \dots, \tau]\}$ 上で考えた $k(x_t, y_t)$ および \mathbb{T} として、以下の汎関数と経路測度を導入する：

$$\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] := \sum_{t'=0}^{t-1} k(x_{t'+1}, y_{t'+1}), \quad \mathbb{P}_F[\mathcal{X}_\tau] := \prod_{t'=0}^{t-1} \mathbb{T}(x_{t'+1}|x_{t'})p(x_0).$$

ここで $p(x_0) = \mathcal{N}_0(x_0) / \sum_{x'} \mathcal{N}_0(x')$ である。これらを用いて、適応度は

$$\Psi[\mathcal{Y}_\tau] := \ln \left\langle e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]} \right\rangle_{\mathbb{P}_F[\mathcal{X}_\tau]}$$

と表される。この表現により、集団適応度が自由エネルギー（キュムラント母関数）と数理的には等価な定義となっていることがわかる。したがって、集団適応度が以下の変分表現を有することも直ちに分かる：

$$\Psi[\mathcal{Y}_\tau] = \max_{\mathbb{P}[\mathcal{X}_\tau]} \left[\left\langle e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]} \right\rangle_{\mathbb{P}[\mathcal{X}_\tau]} - \mathcal{D}_{KL} [\mathbb{P} || \mathbb{P}_F] \right] = \left\langle e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]} \right\rangle_{\mathbb{P}_B[\mathcal{X}_\tau | \mathcal{Y}_\tau]} - \mathcal{D}_{KL} [\mathbb{P}_B || \mathbb{P}_F].$$

³一般に $e^{k_{t+1}(x')} \sum_x \mathbb{T}_{t+1}(x'|x)\mathcal{N}_t(x) \neq \sum_x \mathbb{T}_{t+1}(x'|x)e^{k_{t+1}(x)}\mathcal{N}_t(x)$ なので、表現型が変わってから増殖するのか、増殖してから表現型が変わるのかで問題は変わってくるように見えるが、この問題は \mathcal{N} をどの時間刻みで定義するかだけの問題なので本質的ではない。

⁴時間平均を取る場合もあるが、ここでは積算量として Ψ を考える。

ここで $\mathcal{D}_{KL} [\mathbb{P}||\mathbb{P}_F] := \sum_{\mathcal{X}_\tau} \mathbb{P}_F[\mathcal{X}_\tau] \ln \frac{\mathbb{P}_F[\mathcal{X}_\tau]}{\mathbb{P}[\mathcal{X}_\tau]}$ はカルバックライブラー情報量 (相対エントロピー) である。この変分を達成する測度 $\mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau]$ は陽に

$$\mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau] = \arg \max_{\mathbb{P}[\mathcal{X}_\tau]} \left[\left\langle e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]} \right\rangle_{\mathbb{P}[\mathcal{X}_\tau]} - \mathcal{D}_{KL} [\mathbb{P}||\mathbb{P}_F] \right] = e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] - \Psi[\mathcal{Y}_\tau]} \mathbb{P}_F[\mathcal{X}_\tau]$$

と表される。この測度を時間前向きな経路 $\mathbb{P}_F[\mathcal{X}_\tau]$ に対する時間逆向き (遡及的) な経路測度と呼ぶ。遡及的な経路測度は、時刻 $t = \tau$ で細胞のサンプリングを行い、遡及的に細胞増殖の系譜を遡ってある表現形経路 \mathcal{X}_τ が得られる確率に対応する観測可能量である⁵。また \mathbb{P}_B の形から、統計物理におけるギブス分布と数理的に等価であることがわかる。このように経路積分表現を介することによって、細胞増殖や進化の問題が統計物理と深く関連していることが明らかになる⁶。

この変分関係から対数増殖率や選択の帰結としての $\mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau]$ は、経路に沿って見た子孫の平均再生産数 $\langle \mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] \rangle_{\mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Y}_\tau]}$ を大きくする、という要請と、 \mathbb{P}_B が KL の意味で \mathbb{P}_F からあまり離れないようにする、という2つの要請のバランスによって決定することがわかる。

もし細胞が環境の観測情報 z_t を活用できる場合には、上記の表現は以下のように自然に拡張される：

$$\Psi[\mathcal{Y}_\tau, \mathcal{Z}_\tau] := \ln \left\langle e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]} \right\rangle_{\mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Z}_\tau]} = \max_{\mathbb{P}[\mathcal{X}_\tau|\mathcal{Z}_\tau]} \left[\left\langle e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]} \right\rangle_{\mathbb{P}[\mathcal{X}_\tau|\mathcal{Z}_\tau]} - \mathcal{D}_{KL} [\mathbb{P}||\mathbb{P}_F] \right]$$

ここで、

$$\mathbb{P}_F[\mathcal{X}_t|\mathcal{Z}_t] := \prod_{t'=0}^{t-1} \mathbb{T}(x_{t'+1}|x_{t'}, z_{t'+1}) p(x_0), \quad \mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau] := e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] - \Psi[\mathcal{Y}_\tau, \mathcal{Z}_\tau]} \mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Z}_\tau]$$

である。集団増殖率 Ψ の重要な特性として、基準点の取り方に関する自由度があることがあげられる。具体的には $\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] \rightarrow \mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] + \mathbb{U}$ の変換に対して、 $\Psi[\mathcal{Y}_\tau] \rightarrow \Psi[\mathcal{Y}_\tau] + \mathbb{U}$ が成り立つ。したがって、集団増殖率自体に一定の生物学的意味があるもの、その値そのものは例えば集団に対する選択の強さなどを直接反映せず、より適切な量の導入が必要となる。

3 適応度の最大化

細胞は自分の表現型を変える戦略 \mathbb{T} を進化の過程で最適化してきていると想像できる。では、最適性をどのようにして評価すれば良いのだろうか？

3.1 乗法エルゴード定理と平均対数増殖率

適応度 $\Psi[\mathcal{Y}_\tau]$ についてその時間平均

$$\psi(\tau) := \frac{1}{\tau} \Psi[\mathcal{Y}_\tau],$$

⁵詳細については [22, 23] を参照。

⁶この関係式は非常に一般に成り立つものであり、細胞増殖系のみならず、統計物理、統計推定、情報理論などおよそ確率に関連する様々な分野で広く活用されているものである。

を考える。 $\psi(\tau)$ は集団の拡大率の長時間平均であり、生態学などでは集団の適応度と同一視されることも多い。乗法エルゴード定理の帰結として、比較的素性の良い \mathcal{Y}_τ のダイナミクスのもと、以下の関係が確率 1 で成り立つことが知られている:

$$\lim_{\tau \rightarrow \infty} \psi(\tau) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \langle \Psi[\mathcal{Y}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau]}.$$

この結果は、細胞の最適表現形戦略の指標として

$$\mathbb{P}_F^\dagger = \arg \max_{\mathbb{P}_F} \langle \Psi[\mathcal{Y}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau]},$$

を考えることの基礎となっている。

3.2 環境情報を用いない最適なランダム戦略

ここでは最適な戦略が満たす特性について一般的に解析する。まず細胞は環境の状態を感知することができない、つまり $\mathbb{P}_F[\mathcal{X}_\tau | \mathcal{Y}_\tau, \mathcal{Z}_\tau] = \mathbb{P}_F[\mathcal{X}_\tau | \mathcal{Y}_\tau]$ の場合の最適戦略

$$\mathbb{P}_F^\dagger = \arg \max_{\mathbb{P}_F} \langle \Psi[\mathcal{Y}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau]},$$

を考える。 \mathbb{P}_F^\dagger 周りでの摂動 $\delta \mathbb{P}_F$ と摂動に対する $\langle \Psi \rangle_{\mathbb{Q}}$ の応答 $\delta \langle \Psi \rangle_{\mathbb{Q}}$ を考えることにより、停留条件 $\delta \langle \Psi \rangle_{\mathbb{Q}} = 0$ から最適戦略に関する整合性条件

$$\mathbb{P}_F^\dagger[\mathcal{X}_\tau] = \mathbb{P}_B^\dagger[\mathcal{X}_\tau] := \langle \mathbb{P}_B^\dagger[\mathcal{X}_\tau | \mathcal{Y}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau]}$$

が得られる⁷。

3.3 情報圧縮としての自然選択過程

最適戦略における $\Psi[\mathcal{Y}_\tau]$ の平均は変分表現を用いることにより、

$$\langle \Psi^\dagger \rangle_{\mathbb{Q}} = \langle \mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] \rangle_{\mathbb{P}_B^\dagger[\mathcal{X}_\tau, \mathcal{Y}_\tau]} - \mathcal{I}_B^\dagger$$

となる。ここで

$$\mathbb{P}_B^\dagger[\mathcal{X}_\tau, \mathcal{Y}_\tau] := \mathbb{P}_B^\dagger[\mathcal{X}_\tau | \mathcal{Y}_\tau] \mathbb{Q}[\mathcal{Y}_\tau], \quad \mathcal{I}_B^\dagger = \sum_{x,y} \mathbb{P}_B^\dagger[\mathcal{X}_\tau, \mathcal{Y}_\tau] \ln \frac{\mathbb{P}_B^\dagger[\mathcal{X}_\tau, \mathcal{Y}_\tau]}{\mathbb{P}_B^\dagger[\mathcal{X}_\tau] \mathbb{P}_B^\dagger[\mathcal{Y}_\tau]}$$

は、 \mathbb{P}_B^\dagger を用いた同時分布とその相互情報量である。最適な平均対数増殖率のこの表現は、選択の影響による増殖率のロスが細胞状態の系譜 \mathcal{X}_τ と環境状態の系譜 \mathcal{Y}_τ との間の \mathbb{P}_B^\dagger が定める情報量に等しいことを示す。またこの関係式から $\langle \Psi^\dagger \rangle_{\mathbb{Q}}$ は、別の変分表現

$$\langle \Psi^\dagger \rangle_{\mathbb{Q}} = \max_{\mathbb{P}[\mathcal{X}_\tau | \mathcal{Y}_\tau]} \left[\langle \mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] \rangle_{\mathbb{P}[\mathcal{X}_\tau | \mathcal{Y}_\tau] \mathbb{Q}[\mathcal{Y}_\tau]} - \mathcal{I}[\mathbb{P}[\mathcal{X}_\tau | \mathcal{Y}_\tau] \mathbb{Q}[\mathcal{Y}_\tau]] \right]$$

⁷摂動は \mathbb{P}_F が確率分布である、という制約がついているので、 $\mathbb{P}_F^\dagger(\mathcal{X}_\tau)$ が確率 0 もしくは 1 になるような \mathcal{X}_τ を持っている場合には注意が必要になる。ここではその問題には踏み込まないで、 $\mathbb{P}_F^\dagger[\mathcal{X}_\tau]$ は定義域上で必ず $\mathbb{P}_F^\dagger[\mathcal{X}_\tau] > 0$ とする。またこの時 \mathbb{P}_F^\dagger を生成するマルコフ過程は一般に非定常なものになる。

も有する。ここで $\mathcal{I}[\mathbb{P}[\mathcal{X}_\tau|\mathcal{Y}_\tau]\mathbb{Q}[\mathcal{Y}_\tau]]$ は同時分布 $\mathbb{P}[\mathcal{X}_\tau|\mathcal{Y}_\tau]\mathbb{Q}[\mathcal{Y}_\tau]$ の相互情報量である。これは最適戦略が \mathcal{Y}_τ から \mathcal{X}_τ への情報圧縮と等価であることを表す。 $\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]$ が圧縮の基準となるユーティリティ関数 (負の Distortion) であり、 \mathcal{I}_B^\dagger はユーティリティ関数を大きくするために有効な情報を表す。 $\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]$ は表現形状態の系譜 \mathcal{X}_τ が環境 \mathcal{Y}_τ の時に生産できる子孫数と関連するため、 \mathcal{I}_B^\dagger は子孫生産に実効的に有用となる環境の情報量としてとらえることが可能となる⁸。

4 適応度の満たすゆらぎ定理

最適戦略の満たす整合性条件 $\langle \mathbb{P}_B^\dagger[\mathcal{X}_\tau|\mathcal{Y}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau]} = \mathbb{P}_F^\dagger[\mathcal{X}_\tau]$ for all \mathcal{X}_τ より、

$$\mathbb{P}_B^\dagger[\mathcal{Y}_\tau|\mathcal{X}_\tau] := \frac{\mathbb{P}_B^\dagger[\mathcal{X}_\tau, \mathcal{Y}_\tau]}{\langle \mathbb{P}_B^\dagger[\mathcal{X}_\tau|\mathcal{Y}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau]}} = \frac{e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] - \Psi^\dagger[\mathcal{Y}_\tau]} \mathbb{P}_F^\dagger[\mathcal{X}_\tau] \mathbb{Q}[\mathcal{Y}_\tau]}{\mathbb{P}_F^\dagger[\mathcal{X}_\tau]} = e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] - \Psi^\dagger[\mathcal{Y}_\tau]} \mathbb{Q}[\mathcal{Y}_\tau],$$

となる。 $\mathbb{P}_F^\dagger[\mathcal{X}_\tau]$ とサポートを共有する任意の $\mathbb{P}_F[\mathcal{X}_\tau]$ について

$$e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]} = e^{\Psi[\mathcal{Y}_\tau]} \frac{\mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau]}{\mathbb{P}_F[\mathcal{X}_\tau]},$$

が成り立つことより、

$$\mathbb{P}_B^\dagger[\mathcal{Y}_\tau|\mathcal{X}_\tau] \mathbb{P}_F[\mathcal{X}_\tau] = e^{\Psi[\mathcal{Y}_\tau] - \Psi^\dagger[\mathcal{Y}_\tau]} \mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau] \mathbb{Q}[\mathcal{Y}_\tau],$$

を得る。 $\Delta\Psi := \Psi^\dagger[\mathcal{Y}_\tau] - \Psi[\mathcal{Y}_\tau]$ とし、両辺 \mathcal{X}_τ に関する和を取ると

$$e^{-\Delta\Psi[\mathcal{Y}_\tau]} = \frac{\langle \mathbb{P}_B^\dagger[\mathcal{Y}_\tau|\mathcal{X}_\tau] \rangle_{\mathbb{P}_F[\mathcal{X}_\tau]}}{\mathbb{Q}[\mathcal{Y}_\tau]},$$

を得る。この関係式を適応度に関する詳細ゆらぎ関係と呼ぶ。両辺を $\mathbb{Q}[\mathcal{Y}_\tau]$ について平均を取ると直ちに

$$\langle e^{-\Delta\Psi[\mathcal{Y}_\tau]} \rangle_{\mathbb{Q}[\mathcal{Y}_\tau]} = 1,$$

がわかる。これを適応度に関する積分ゆらぎ関係と呼ぶ。また詳細ゆらぎ関係の両辺を対数をとってから $\mathbb{Q}[\mathcal{Y}_\tau]$ について平均を取ると

$$\langle \Psi^\dagger \rangle_{\mathbb{Q}} - \langle \Psi \rangle_{\mathbb{Q}} = \langle \Delta\Psi[\mathcal{Y}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau]} = \mathcal{D}_{KL} \left[\mathbb{Q}[\mathcal{Y}_\tau] \parallel \langle \mathbb{T}_B^\dagger[\mathcal{Y}_\tau|\mathcal{X}_\tau] \rangle_{\mathbb{P}_F[\mathcal{X}_\tau]} \right],$$

を得る。これを平均ゆらぎ関係と呼ぶ。平均ゆらぎ関係は準最適な \mathbb{P}_F と最適な \mathbb{P}_F^\dagger を比較した時、平均対数増殖率の損失が KL 距離で定まることを示す。また、積分ゆらぎ関係は $\Delta\Psi[\mathcal{Y}_\tau]$ のゆらぎがある種の恒等式を満たすことを示す。そして詳細ゆらぎ関係は、ある環境における $\Delta\Psi[\mathcal{Y}_\tau]$ が環境のゆらぎ \mathbb{Q} やその環境における最適戦略によって定まる $\mathbb{P}_B^\dagger[\mathcal{Y}_\tau|\mathcal{X}_\tau]$ とどのように関連しているかを明示している。またゆらぎ関係が Ψ ではなく $\Delta\Psi$ について成り立つ理由は、 Ψ が基準点の取り方に関する自由度に起因する。その意味で、 Ψ^\dagger は様々な状況における適応度を測る基準点としての役割をはたすこともゆらぎ関係からわかる。

⁸例えば環境 \mathcal{Y}_τ と環境 \mathcal{Y}'_τ の元で細胞は自身の状態によらず全く同じ増殖率を持つ、つまり $\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] = \mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}'_\tau]$ とする。この時、環境変動 \mathcal{Y}_τ と \mathcal{Y}'_τ を区別する情報は増殖には全く関係せず、したがって子孫生産には有効でない情報になる。

5 環境情報を用いた最適なランダム戦略

上記の議論は細胞が環境の観測によって情報 \mathcal{Z}_τ を得られる場合についても自然に拡張される。細胞は環境の状態を感知することができる、つまり $\mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Z}_\tau]$ の場合を考える。ここで、環境と観測の同時分布を $\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]$ とするとき、最適戦略

$$\mathbb{P}_F^* = \arg \max_{\mathbb{P}_F} \langle \Psi[\mathcal{Y}_\tau, \mathcal{Z}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]}$$

を考える。環境情報を用いない場合の最適戦略と区別するために \dagger ではなく $*$ を用いている。 \mathbb{P}_F^* 周りでの摂動 $\delta\mathbb{P}_F$ と摂動に対する $\langle \Psi \rangle_{\mathbb{Q}}$ の応答の停留条件 $\delta \langle \Psi \rangle_{\mathbb{Q}} = 0$ から、整合性条件

$$\langle \mathbb{P}_B^*[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau|\mathcal{Z}_\tau]} = \mathbb{P}_F^*[\mathcal{X}_\tau|\mathcal{Z}_\tau]$$

が得られる⁹。

5.1 適応度と情報の満たすゆらぎ定理

最適戦略の満たす停留条件 $\langle \mathbb{P}_B^*[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau|\mathcal{Z}_\tau]} = \mathbb{P}_F^*[\mathcal{X}_\tau|\mathcal{Z}_\tau]$ for all \mathcal{X}_τ and \mathcal{Z}_τ より、

$$\mathbb{P}_B^*[\mathcal{Y}_\tau|\mathcal{X}_\tau, \mathcal{Z}_\tau] := \frac{\mathbb{P}_B^*[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau]\mathbb{Q}[\mathcal{Y}_\tau|\mathcal{Z}_\tau]}{\langle \mathbb{P}_B^*[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau|\mathcal{Z}_\tau]}} = \frac{\mathbb{P}_B^*[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau]\mathbb{Q}[\mathcal{Y}_\tau|\mathcal{Z}_\tau]}{\mathbb{P}_F^*[\mathcal{X}_\tau|\mathcal{Z}_\tau]} = e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau] - \Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau]}\mathbb{Q}[\mathcal{Y}_\tau|\mathcal{Z}_\tau]$$

よって、

$$\mathbb{P}_B^*[\mathcal{X}_\tau, \mathcal{Y}_\tau, \mathcal{Z}_\tau] := \mathbb{P}_B^*[\mathcal{Y}_\tau|\mathcal{X}_\tau, \mathcal{Z}_\tau]\mathbb{P}_F^*[\mathcal{X}_\tau|\mathcal{Z}_\tau]\mathbb{Q}[\mathcal{Z}_\tau] = \mathbb{P}_B^*[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau]\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]$$

が成り立つ。また $\mathbb{P}_F^*[\mathcal{X}_\tau|\mathcal{Z}_\tau]$ とサポートを共有する任意の $\mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Z}_\tau]$ について、 $e^{\mathbb{K}[\mathcal{X}_\tau, \mathcal{Y}_\tau]} = e^{\Psi[\mathcal{Y}_\tau, \mathcal{Z}_\tau]}\mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau]/\mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Z}_\tau]$ が成り立つことから、

$$\mathbb{P}_B^*[\mathcal{Y}_\tau|\mathcal{X}_\tau, \mathcal{Z}_\tau]\mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Z}_\tau]\mathbb{Q}[\mathcal{Z}_\tau] = e^{\Psi[\mathcal{Y}_\tau, \mathcal{Z}_\tau] - \Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau]}\mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau]\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]$$

が成り立つ。 $\Delta\Psi^* := \Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau] - \Psi[\mathcal{Y}_\tau, \mathcal{Z}_\tau]$ とすると

$$e^{-\Delta\Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau]} = \frac{\mathbb{P}_B^*[\mathcal{Y}_\tau|\mathcal{X}_\tau, \mathcal{Z}_\tau]\mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Z}_\tau]}{\mathbb{P}_B[\mathcal{X}_\tau|\mathcal{Y}_\tau, \mathcal{Z}_\tau]\mathbb{Q}[\mathcal{Y}_\tau|\mathcal{Z}_\tau]}$$

を得る。ここで $\mathbb{P}_F[\mathcal{X}_\tau|\mathcal{Z}_\tau]$ として環境情報を使わない場合の最適戦略 $\mathbb{P}_F^\dagger[\mathcal{X}_\tau]$ を用いると

$$e^{-[\Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau] - \Psi^\dagger[\mathcal{Y}_\tau]]} = \frac{\mathbb{P}_B^*[\mathcal{Y}_\tau|\mathcal{X}_\tau, \mathcal{Z}_\tau]\mathbb{P}_F^\dagger[\mathcal{X}_\tau]}{\mathbb{P}_B^\dagger[\mathcal{X}_\tau|\mathcal{Y}_\tau]\mathbb{Q}[\mathcal{Y}_\tau|\mathcal{Z}_\tau]} = \frac{\mathbb{P}_B^*[\mathcal{Y}_\tau|\mathcal{X}_\tau, \mathcal{Z}_\tau]\mathbb{P}_F^\dagger[\mathcal{X}_\tau]}{\mathbb{P}_B^\dagger[\mathcal{X}_\tau|\mathcal{Y}_\tau]\mathbb{Q}[\mathcal{Y}_\tau]} \frac{\mathbb{Q}[\mathcal{Y}_\tau]\mathbb{Q}[\mathcal{Z}_\tau]}{\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]}$$

今、 $i[\mathcal{Y}_\tau, \mathcal{Z}_\tau] := \ln \frac{\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]}{\mathbb{Q}[\mathcal{Y}_\tau]\mathbb{Q}[\mathcal{Z}_\tau]}$ とすると、適応度と情報と詳細ゆらぎ関係

$$e^{\Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau] - \Psi^\dagger[\mathcal{Y}_\tau] - i[\mathcal{Y}_\tau, \mathcal{Z}_\tau]} = \frac{\mathbb{P}_B^\dagger[\mathcal{X}_\tau|\mathcal{Y}_\tau]\mathbb{Q}[\mathcal{Y}_\tau]}{\mathbb{P}_B^*[\mathcal{Y}_\tau|\mathcal{X}_\tau, \mathcal{Z}_\tau]\mathbb{P}_F^\dagger[\mathcal{X}_\tau]} = \frac{\mathbb{P}_B^\dagger[\mathcal{Y}_\tau|\mathcal{X}_\tau]}{\mathbb{P}_B^*[\mathcal{Y}_\tau|\mathcal{X}_\tau, \mathcal{Z}_\tau]} = \frac{\mathbb{P}_B^\dagger[\mathcal{Y}_\tau|\mathcal{X}_\tau]\mathbb{P}_F^*[\mathcal{X}_\tau|\mathcal{Z}_\tau]\mathbb{Q}[\mathcal{Z}_\tau]}{\mathbb{P}_B^*[\mathcal{X}_\tau, \mathcal{Y}_\tau, \mathcal{Z}_\tau]}$$

⁹基本的には \mathcal{Z}_τ での条件付けが入っただけである。ただし $\mathbb{P}_F^*[\mathcal{X}_\tau|\mathcal{Z}_\tau]$ は一般には因果性を満たさないため、生物学的な意味を持つ状況は限定され、数理的な意味合いが強くなる。

をえる。両辺 $\mathbb{P}_B^*[\mathcal{X}_\tau, \mathcal{Y}_\tau, \mathcal{Z}_\tau]$ で和を取ると

$$\left\langle e^{\Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau] - \Psi^\dagger[\mathcal{Y}_\tau] - i[\mathcal{Y}_\tau, \mathcal{Z}_\tau]} \right\rangle_{\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]} = 1$$

をえる。また対数を取り両辺 $\mathbb{P}_B^*[\mathcal{X}_\tau, \mathcal{Y}_\tau, \mathcal{Z}_\tau]$ で和を取ると

$$\left\langle \Psi^\dagger[\mathcal{Y}_\tau] \right\rangle_{\mathbb{Q}[\mathcal{Y}_\tau]} + \mathcal{I} - \langle \Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]} = \mathcal{D}_{KL} \left[\mathbb{P}_B^*[\mathcal{X}_\tau, \mathcal{Y}_\tau, \mathcal{Z}_\tau] \parallel \mathbb{P}_B^\dagger[\mathcal{Y}_\tau | \mathcal{X}_\tau] \mathbb{P}_F^*[\mathcal{X}_\tau | \mathcal{Z}_\tau] \mathbb{Q}[\mathcal{Z}_\tau] \right] \geq 0$$

をえる。ここで $\mathcal{I} := \sum_{\mathcal{Y}_\tau, \mathcal{Z}_\tau} \mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau] \ln \frac{\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]}{\mathbb{Q}[\mathcal{Y}_\tau] \mathbb{Q}[\mathcal{Z}_\tau]}$ は環境と環境情報の間の相互情報量である。この関係から直ちに

$$\left\langle \Psi^\dagger[\mathcal{Y}_\tau] \right\rangle_{\mathbb{Q}[\mathcal{Y}_\tau]} + \mathcal{I} \geq \langle \Psi^*[\mathcal{Y}_\tau, \mathcal{Z}_\tau] \rangle_{\mathbb{Q}[\mathcal{Y}_\tau, \mathcal{Z}_\tau]}$$

となる。この関係式は環境についての情報 \mathcal{Z}_τ を使った場合の適応度の利得は、使わなかった場合の最適な戦略と比較して、 \mathcal{Z}_τ の持つ環境についての情報量 \mathcal{I} でおさえられていることを示す。つまり、 \mathcal{Z}_τ の適応的な価値は最大でもその情報量 \mathcal{I} であることが示される。しかしながら情報量 \mathcal{I} は必ずしも不等式のタイトなバウンドではない。例えば、 \mathcal{Z}_τ が増殖とは無関係の環境状態 \mathcal{Y}_τ についての情報を多量に持っていた場合、それは \mathcal{I} を大きくはするが、適応度の増加には寄与しない。また統合を達成する条件も明確ではない。よりタイトな bound は \mathcal{X}_τ と \mathcal{Y}_τ 、そして \mathcal{Z}_τ の間の3体の相互情報量によって決定する¹⁰。

6 まとめ

以上の結果から、細胞集団などの増殖系の振る舞いは統計物理との関連が極めて強いことが明らかになった。そこで重要な役割をはたすのが \mathbb{P}_B や \mathcal{I}_B などに表現される遡及的な現象見方である。遡及的な見方とは、現象をある時刻から遡って事後的に見ることであり、したがって増殖系における選択の本質は事後的な情報処理である。一方、環境を細胞が観測してその振る舞いを変える $\mathbb{P}_F[\mathcal{X}_\tau | \mathcal{Z}_\tau]$ はその生物学的な仕組みから事前に行われる情報処理である。 \mathbb{P}_F はそのため因果性を持つという拘束の中で進化的に変化をすることになる。これに対し、自然選択による遡及的な \mathbb{P}_B は因果性を持つ経路測度では無い。つまり、 \mathbb{P}_B で実現可能な振る舞いと \mathbb{P}_F で実現可能な振る舞いは本質的に異なる。この2つの違いが進化的に環境に事後的に適応する選択の過程と、環境を観測しその情報をもとに事前に適応する情報処理との違いである。経路測度を用いた増殖系の表現はこのような適応の2つの側面を統合的に表現し解析することを可能とする。今後この表現を用いることにより、生物適応のより本質的な理解につながることを期待される。

参考文献

- [1] K A Geiler-Samerotte, C R Bauer, S Li, N Ziv, D Gresham, and M L Siegal. The details in the distributions: why and how to study phenotypic variability. *Curr. Opin. Biotech.*, 24(4):752–759, August 2013.

¹⁰導出は複雑なので [23] を参照。

- [2] Nathalie Q Balaban, Jack Merrin, Remy Chait, Lukasz Kowalik, and Stanislas Leibler. Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625, September 2004.
- [3] Y Wakamoto, N Dhar, R Chait, K Schneider, F Signorino-Gelo, S Leibler, and J D McKinney. Dynamic Persistence of Antibiotic-Stressed Mycobacteria. *Science*, 339(6115):91–95, January 2013.
- [4] R Jayaraman. Bacterial persistence: some new insights into an old phenomenon. *Journal of Biosciences*, 33(5):795–805, 2008.
- [5] Neeraj Dhar and J D McKinney. Microbial phenotypic heterogeneity and antibiotic tolerance. *Current Opinion in Microbiology*, 10(1):30–38, 2007.
- [6] Cancer therapy resistance: chasing epigenetics. *Nat. Med.*, 20(4):340–341, April 2014.
- [7] Corbin E Meacham and Sean J Morrison. Tumour heterogeneity and cancer cell plasticity. *Nature*, 501(7467):328–337, September 2013.
- [8] Richard Levins. *Evolution in Changing Environments*. Some Theoretical Explorations. Princeton University Press, 1968.
- [9] Patsy Haccou, Sake J De Vlas, Jacques JM Van Alphen, and Marcel E Visser. Information processing by foragers: effects of intra-patch experience on the leaving tendency of *Leptopilina heterotoma*. *The Journal of Animal Ecology*, pages 93–106, 1991.
- [10] P Haccou and Y Iwasa. Establishment probability in fluctuating environments: a branching process model. *Theoretical Population Biology*, 50(3):254–280, December 1996.
- [11] P Haccou and Y Iwasa. Optimal mixed strategies in stochastic environments. *Theoretical Population Biology*, 47(2):212–243, 1995.
- [12] Patsy Haccou and Yoh Iwasa. Robustness of optimal mixed strategies. *J. Math. Biol.*, 36(5):485–496, 1998.
- [13] Imke G de Jong, Patsy Haccou, and Oscar P Kuipers. Bet hedging or not? A guide to proper classification of microbial survival strategies. *Bioessays*, 33(3):215–223, 2011.
- [14] Edo Kussell and Stanislas Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075–2078, 2005.
- [15] Eshel Ben-Jacob and Daniel Schultz. Bacteria determine fate by playing dice with controlled odds. *Proc. Natl. Acad. Sci. U.S.A.*, 107(30):13197–13198, 2010.
- [16] *Shannon information and biological fitness*. IEEE, 2004.
- [17] Matina C Donaldson-Matasci, Carl T Bergstrom, and Michael Lachmann. The fitness value of information. *Oikos*, 119(2):219–230, February 2010.
- [18] Tetsuya J Kobayashi and Atsushi Kamimura. Theoretical aspects of cellular decision-making and information-processing. *Adv. Exp. Med. Biol.*, 736:275–291, 2012.
- [19] Olivier Rivoire and Stanislas Leibler. The value of information for populations in varying environments. *J Stat Phys*, 142(6):1124–1166, March 2011.

- [20] Olivier Rivoire and Stanislas Leibler. A model for the generation and transmission of variations in evolution. *Proc. Natl. Acad. Sci. U.S.A.*, 111(19):E1940–9, May 2014.
- [21] Edo Kussell. Evolution in Microbes. *Annu. Rev. Biophys.*, 42(1):493–514, May 2013.
- [22] Yuki Sughiyama, Tetsuya J Kobayashi, Koji Tsumura, and Kazuyuki Aihara. Pathwise thermodynamic structure in population dynamics. *Phys Rev E Stat Nonlin Soft Matter Phys*, 91(3):032120–9, March 2015.
- [23] Tetsuya J Kobayashi and Yuki Sughiyama. Fluctuation Relations of Fitness and Information in Population Dynamics. *arXiv*, May 2015.