

遺伝子進化と確率過程

国立遺伝学研究所

太田朋子

1 はじめに

生物の進化はいうまでもなくC.ダーウィンの自然淘汰説によって説明されている。ダーウィンが種の起源を出版したのは1859年であるが、出版によって進化論の一大論争が起ったことはよく知られている。注目すべきは、ダーウィンに反対するものの中に自然淘汰を排除して環境の影響が進化の主因であると主張する一派が現われた。これが新ラマルタ主義である。これに真正面から反対したのがA. ヴァイスマンでマウスの尾を19世代も切っても尾の短いマウスは得られないという実験で示したのも有名である。

一方、G.メンデルはダーウィンと同時期に修道院の庭でエンドウマメの交雑実験を行っていた。メンデルは世代を通して遺伝的形質を伝える因子（後に遺伝子と呼ばれる）があることをつきとめ雑種植物の研究という論文を1866年に発表した。これが確率・統計的考えに基づいていたこともあり、当時その真価が認められなかった。19世紀末にイギリスには生物統計学の研究が始まり、量的形質がまじりあってしまうので、自然淘汰が働けなくなってしまうという融合遺伝の考えにたって、激しい論争が繰り返された。

1900年にメンデル法則の再発見により今世紀にメンデル遺伝学は大発展をとげた。そこで自然淘汰説とメンデル遺伝学とを、統計学の方法を用いて結合しようとしてきたのが集団遺伝学である。1930年にR.A.フィッシャー、J.B.S.ホールデン、S.ライトの3人によってその基礎が作られた。ここで量的形質もいくつかの遺伝子が関与していて、メンデル遺伝で分離するとすれば、融合遺伝の考えも間違っていたことがはっきりした。自然淘汰説で充分理解できるとされて、ネオ・ダーウィニズム、または新総合説の全盛時代となった。一方、化学や物理学の原理で生命現象を理解しようとする動きが高まってきた。

遺伝物資はどういうものか。E.シュレージンガーの『生命とは何か』という本は生物と物理を結ぶ上で、多大な影響を与えた。遺伝子は 1) 情報を持つ 2) 自己複製する 3) 突然変異を起すという性質を持つといった根本原理が論議された。これらの性質を持つ遺伝子として申し分のない物質がDNAである。ワトソン-クリックモデルは1953年発表された。それで生命科学は変革期を迎えた。M.デルブルックなどの物理学者もファージなどを用いて遺伝学に真剣に取り組んだ。分子遺伝学の全盛期に入るわけである。しかし分子レベルで進化を論ずることができるようになったのは1960年代の後半である。蛋白質の一次構造の変化から生物の進化が考察されたのは、ヘモグロビンがその最初である。E.ズッカーカンドルとL.ポーリングはヘモグロビンとチトクロームcのアミノ酸配列の比較から1965年に分子進化の速度はかなり一定らしい、すなわち分子進化時計の考え方を提唱した。そして集団遺伝学と分子進化とを結合して木村資生は分子進化中立説を1968年に提唱した。これは分子進化が自然淘汰によくも悪くもない中立な突然変異が偶然に置きかわった結果であると主張するもので、ゲノム全体で見ると、突然変異遺伝子の置換の数が、ホールデンの予想をはるかに越えてしまうという事実に基づいて提唱された。この説が当時のネオ・ダーウィニズムとは相いれない説であったため大きな論争となった。

私は単純化された中立理論に満足できず、中立突然変異と自然淘汰の効果と共に考慮した遺伝子進化の理論、すなわちほぼ中立理論について研究してきた。現在では遺伝子データの解析による理論の検証も行っている。以下に私がほぼ中立理論を提唱するに至った状況とその発展の過程について述べる。

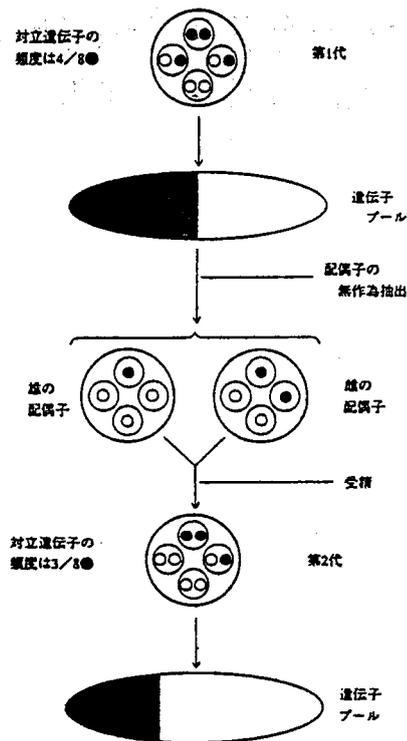


Figure 1: 遺伝子頻度の偶然的变化を模式的に示した図。4 個体からなる仮想集団で●遺伝子の頻度が4/8から3/8に変化する（文献3）。

2 集団中における突然変異遺伝子の行動

実際の進化の過程では、生物集団中には常に新しい突然変異が生じており、その多くは有害で、数代の後に集団から自然淘汰により除去される。しかし有害でない、すなわち自然淘汰に無関係（中立）またはごくまれには有利な突然変異が生じており、これらが進化の原動力となっているわけである。注意すべき点は、中立な突然変異はいうに及ばず、有利なものでも、たった1個だけ集団に起った突然変異についていえば、大部分が集団の中から消失する点である。その理由は、生物は繁殖に際して自己のもっているあらゆる遺伝子を子供に伝えるわけにはいかないから。生物集団は多くの場合、その大きさが代々ほとんど一定に保たれ無制限に大きくなることはないわけで、ある遺伝子が親から子へ伝えられるのは自然淘汰に有利であっても偶然に依存する。

このような遺伝子頻度の繁殖にともなう偶然による変動は遺伝的浮動と呼ばれ、一般に誤解されやすいので図1を用いて次に説明しよう。いま4個体からなり第1代では、 A_1 (図の●) と A_2 (○) の遺伝子頻度が50%であるような仮想的な集団を考えよう。1世代で沢山の配偶子が生じ、 A_1 、 A_2 、50%ずつからなる遺伝子の集合（遺伝子プールと呼ぶ）ができるが、生殖時にはほんのわずかが偶然的に抽出される。このため図に示すように A_1 、 A_2 の頻度がランダムに増減する。

さて遺伝子頻度が偶然の影響と、方向の定まった自然淘汰の影響の両方のもとに変化するとき、どちらがより有効に働くかが問題となる。いま集団の次代に寄与する親の個体数を N とすると、親の配偶子は $2N$ 個の集団ということになる。ある遺伝子座に注目すれば、 $2N$ 個の相同遺伝子を抽出することになり、それによってどれ程遺伝子頻度がゆらぐかということである。自然淘汰に s だけ有利な突然変異遺伝子は、他の対立遺伝子に比べ s だけ余分に次代に伝わっていく。

3 拡散モデルの適用

このような遺伝子頻度の変化に関する確率過程を取り扱うために物理学の拡散モデルを適用できる。木村資生はフィッシャーやライトのこうした試みを大きく発展させた（文献2）。以下にその方法の概略を説明する。

N個体からなる集団にAとaの対立遺伝子がそれぞれ頻度xとyで分離しているとする（x+y=1）。xの変化に関する確率分布を次のようなコルモゴロフ方程式（フォッカー・プラック方程式ともいう）で書き表わすことができる。

$$\frac{\partial \phi}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial x^2} \{V_{\delta x} \phi\} - \frac{\partial}{\partial x} \{M_{\delta x} \phi\} \quad (1)$$

ここで $\phi = \phi(p, x; t)$ はxの頻度が時間t=0のときpであったとしてt世代後にxになる確率密度を表わす。また $M_{\delta x}$ と $V_{\delta x}$ はそれぞれ1世代あたりのxの変化の平均と分散である。（1）式はコルモゴロフの上向方程式と呼ばれ、この式を用いていろいろな場合の遺伝子頻度の確率分布を求めることができる。

次にxを固定し、pを変数とみなせばコルモゴロフの下向方程式となる。

$$\frac{\partial \phi(p, x; t)}{\partial t} = \frac{V_{\delta p}}{2} \frac{\partial^2 \phi(p, x; t)}{\partial p^2} + M_{\delta p} \frac{\partial \phi(p, x; t)}{\partial p} \quad (2)$$

ここでx=1とおけば、 ϕ は初期頻度pであった対立遺伝子Aがt世代までに集団の中に拡がって固定する割合となる。これをu(p, t)とおいて、

$$\frac{\partial u(p, t)}{\partial t} = \frac{V_{\delta p}}{2} \frac{\partial^2 u(p, t)}{\partial p^2} + M_{\delta p} \frac{\partial u(p, t)}{\partial p} \quad (3)$$

が固定確率を与える式である。遺伝子進化で興味深いのは究極的固定確率、 $u(p) = \lim_{t \rightarrow \infty} u(p, t)$ である。これは次の微分方程式を満足する。

$$\frac{V_{\delta p}}{2} \frac{d^2 u(p)}{dp^2} + M_{\delta p} \frac{du(p)}{dp} = 0 \quad (4)$$

ここでU(0)=0, U(1)=1。この式を用いて有限集団中に突然変異遺伝子が固定する確率を計算することができる。

図1のような有限集団のサンプリングによる遺伝子頻度の変化の分散は $V_{\delta p} = p(1-p)/(2N)$ となる。また自然淘汰のモデルで一番簡単な場合は、AA, Aa, aaの適応度をそれぞれ1+s, 1+(s/2), 1とし、ヘテロ接合体（Aa）の値を両ホモ接合体の中間にもっていった場合である。ここでsは淘汰係数である。この時1世代の遺伝子頻度の変化の平均値は、おおよそ $M_{\delta p} = (s/2)p(1-p)$ となる。これらの値を用い（4）式をとけば固定確率の式が得られる。

$$u(p) = \int_0^p G(x) dx / \int_0^1 G(x) dx \quad (5)$$

ここで

$$G(x) = \exp \left\{ - \int \frac{M_{\delta x}}{V_{\delta x}} dx \right\}.$$

図2は計算結果を示した図である。すなわち突然変異の固定確率が淘汰の強さとどう関係するかをあらわし、横軸は集団の大きさと淘汰係数との積で縦軸は固定確率である。連続関数であるが、注意しなければならないのは淘汰係数が負の鎮域すなわち有害淘汰の場合である。この場合Nsの絶対値が

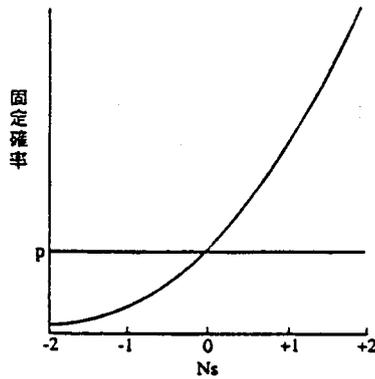


Figure 2: 突然変異遺伝子の固定確率を N_s の関数として表わした図。 p は初期頻度を表わす (文献 4)。

1より充分大きければ、自然淘汰の影響のほうが強くて、そのような不利な突然変異遺伝子は常に集団から自然淘汰により除去されて、広がっていくことはできない。しかし N_s の絶対値が1より小さいような場合は、 s が負であってもごくわずかではあるがひろがり得るわけで、次節で述べるように分子進化を理解する上で重要と考えられる。このとき N_s の絶対値とひろがっていく割合との間には負の相関があり、有害な度合 s が大きければ大きい程、また集団の大きさ N が大きければ大きい程、広がっていないわけである。これに対し s が正の領域すなわち一般にいうダーウィン淘汰では N_s が大きい程固定確率は高い。したがってもし s が一定であれば、集団が大きい程固定確率が高いと予測できる。以上は s を一定とし単純化した考え方であるが、より現実的に s についてゼロ近傍の分布を仮定し、集団全体の適応度の増減を取り入れたモデルでも、ほぼ中立な突然変異については集団の大きさと進化速度との間には負の相関が存在することがわかった。この場合ほぼ中立という意味は弱有害のみでなく弱有利も含み遺伝子進化一般を理解するのに都合がよい。

4 遺伝子進化

まず最もよく研究の進んでいるヘモグロビンの遺伝子進化について述べよう。高等な脊椎動物ではヘモグロビンの分子は α 鎖2本と β 鎖2本からなるが、そのうち α 鎖は哺乳類では141のアミノ酸がつながってできたものである。この141のアミノ酸配列を各種生物について比較してみると、違いの程度は生物の類縁関係と平行になっていることがわかる。次の図3はいくつかの生物の系統樹と、ヘモグロビン α 鎖を比較したときのアミノ酸の違いの数が示してある。一方古生物学の研究から、ウマとヒトは今から約8000万年前に分かれたことがわかっているので、進化の過程である一つのアミノ酸座位に置換の起る率は現時点でウマとヒトは18個のアミノ酸が異なっているから平均して1年あたり

$$(18/141) \div (8 \times 10^7) \div 2 = 0.8 \times 10^{-9}$$

となる。これはほぼ10億年に1個の変化に相当する。また2で割るのは共通の祖先からヒトとウマへの二つの経路を取って進化が起ってきたからである。

こういった比較をいくつか取り計算してみると、興味深いことにヘモグロビン α の遺伝子はどの生物の系統でもほとんど一定の速度で進化してきていることが分かる。図3は各種生物のヘモグロビン α の系統樹を表わす。このことから分子進化は表現型のレベルでの進化とは無関係で、生きた化石といわれるような種と、表現型の上では速く進化している種とで、ヘモグロビン遺伝子の一次構造については同じように進化が起ってきていることが分かる。こうした事実は従来の進化学説からはまったく予想されなかったわけで、中立説でのみ理解できる。中立突然変異についていえば、進化の速度と

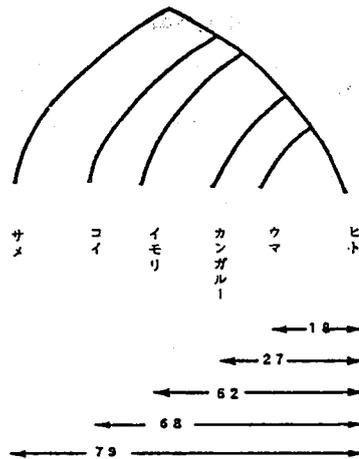


Figure 3: ヘモグロビン α の一次構造が各種生物の間でどの程度違っているかを示し、アミノ酸配列を比較したときの異なったアミノ酸の数で表わしてある。

突然変異率が等しく環境の変化や集団の大きさに依存しないのだから。しかし現実の進化過程はそれ程単純ではない。まず第一は蛋白質の進化速度は各々の蛋白質で異なり、機能的制約に応じてさまざまな速度で進化する点である。表1はいくつかの例を示す。中立説では突然変異率はそれ程違わないが、出てくる突然変異の中で機能をそこなう、したがって有害効果をもつものの割合が異なるためと説明される。

タンパク質	
フィブリノペプチド	8.3
膵臓リボヌクレアーゼ	2.1
リゾチーム	2.0
ヘモグロビン α	1.2
ミオプロビン	0.89
インシュリン	0.44
チトクロームc	0.3
ヒストン H4	0.01

Table 1: アミノ酸置換で表わしたタンパク質の進化速度。これらの速度は、ほとんどが哺乳類のデータに基づいており、年当り 10^{-9} を単位としてアミノ酸座位当たりで表わしてある（文献3）。

ここで有害とそうでないような突然変異の境界はどうかという問題が生ずる。自然淘汰が全か無かという程単純に働くとは考えにくいので、私は境界としてのほぼ中立な突然変異が大切であると提唱した。図4は単純な中立説とほぼ中立理論とを比較した模式図である。

先に説明した固定確率との関連のもとに、ほぼ中立な突然変異による進化について考えてみる。もし自然淘汰に有利な突然変異がある一定の突然変異率で集団中に現れるとすれば、進化速度は集団の大きさと正の相関があるだろう。もし完全に中立であるならば、進化速度は突然変異率に等しい。ここで問題となるのは、突然変異率は、経験的に世代あたりおおよそ一定であるといわれているのに、進化の速度はヘモグロビン α の場合からも明らかのように、なぜ年あたり一定になるかという問題である。これに対しては、突然変異の多くが完全に中立ではなく、ごくわずかに有害効果をもっていると考えたほうが合理的である。なぜならそのような突然変異では進化速度は集団の大きさと負の相関

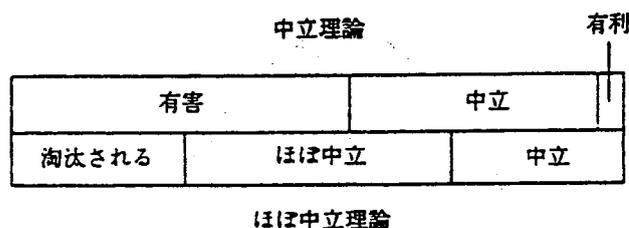


Figure 4: 新しく生ずる突然変異に関する中立理論とほぼ中立理論を比較する模式図 (文献4)

Star phylogenies of 49 genes

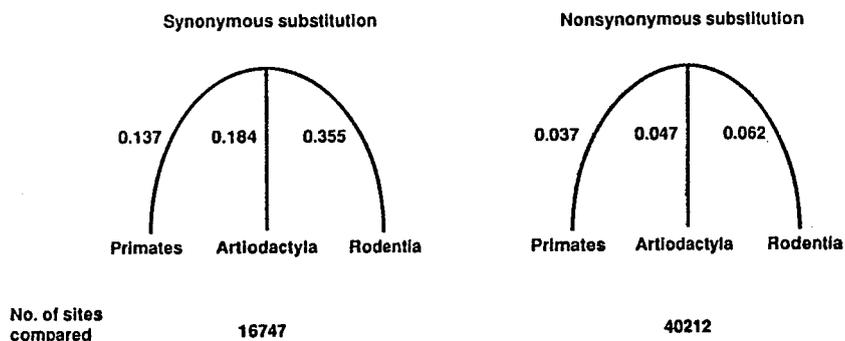


Figure 5: 49の遺伝子における哺乳類系統樹の同義置換および非同義置換の総数の推定値

が予想されるが、一般に集団の小さい生物では体が大きく世代が長い事実と考えあわせれば、進化速度が1年あたりほぼ一定になるのを説明しやすいから。この考えはタンパク質の高次構造と機能に関する制約という面からも理解しやすい。

私は20年程前ほぼ中立理論に基づいて次のような予測をたてた。その当時ヒトのゲノムに存在する30億ものDNA塩基対の中で、真に遺伝情報として働いている部分はほんの数パーセントにすぎないだろうと推論されていた。とすればゲノムDNAの大部分については、その塩基配列が生物個体の生存に関係しないから、塩基置換が完全に中立な突然変異としての行動をとるであろう。これに対しアミノ酸置換は先に述べたほぼ中立な突然変異として行動するのではなかろうか。したがってゲノムDNAの大部分の進化速度は集団の大きさに依存せずもっぱら突然変異率によるのに対し、アミノ酸置換速度は集団の大きさとは負の相関があるであろう。哺乳類についてみると、一般にネズミのように体の小さい動物は集団サイズが大きく世代が短いに対し、大型のサルでは集団サイズが小さく世代が長い。したがってアミノ酸進化速度とDNA進化速度との比をとればネズミの方がサルやヒトより低くなるはずである。

現在ではこの予測をDNA塩基配列の比較により検定することができる。49の遺伝子の塩基配列をヒト、ネズミ、およびウシの類で比較、アミノ酸に変化を起さないいわゆる同義置換とアミノ酸の変化を生ずる非同義置換とにわけて進化距離の推定をした。さらにこれらの系統が分岐して以来の枝の長さをもとめ図示したのが図5である。同義置換がネズミの系統で速いという傾向がはっきりみられる。同義置換は中立と考えられるから、これに対しネズミでアミノ酸の変化を起す置換が少ないということは先程の予測とうまくあっている。ネズミは体が小さく霊長類に比べて集団が大きいから、ネズミの方が自然淘汰が有効に働きほぼ中立なアミノ酸置換が起りにくいわけである。

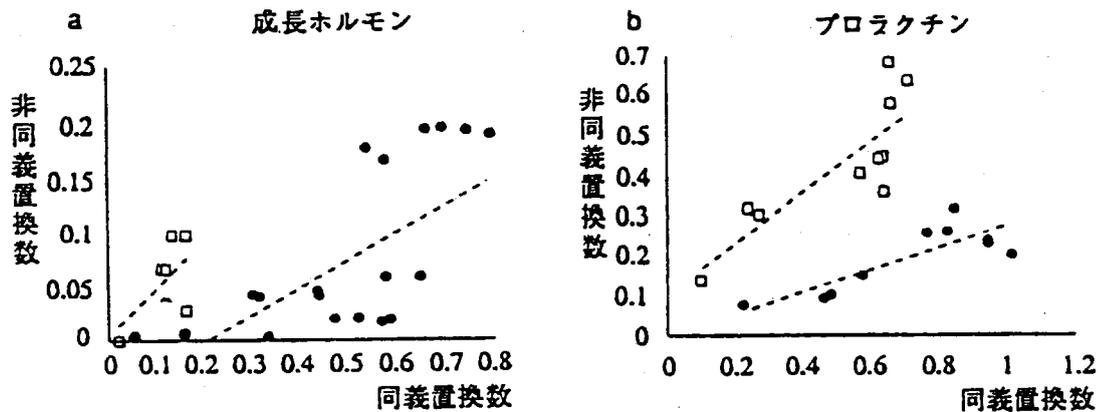


Figure 6: 成長ホルモンおよびプロラクチンの遺伝子比較による座位あたりの同義置換数と非同義置換数の関係を示す。□は重複遺伝子の比較で●は種間遺伝子の比較である。点線は回帰直線を表わす。

5 DNAの柔軟化と進化

これまで考えてきたことは1つの固定した遺伝子座についていえることである。しかし高等生物のゲノムは複雑で、遺伝子が分断されていたり、重複した遺伝子群や部分交換をしたと思われるモザイク遺伝子が普遍的に存在する。すなわち染色体の細かい重複や欠失、組換えといった変化が進化的に重要であることが示される。DNAは従来考えられていたよりも柔軟性にとんでいるといえそうである。私はDNAの重複や欠失を考慮したモデルで遺伝子族の進化について研究してきたが、その結果新しい機能をもった遺伝子族ができてくるためには、機能をふやすような積極的な意味の自然淘汰の働く必要性を示すことができた。

さて現在では各種遺伝子族のデータ解析によりこの理論を検証することができる。私は生長ホルモンの遺伝子族を調べてみた。生長ホルモンとプロラクチンは相同性が高く同じ遺伝子族に属する。多くの哺乳類は生長ホルモンの遺伝子1つとプロラクチンの遺伝子1つを持つが、ヒトでは生長ホルモンの遺伝子が5つ並んで存在している。またウシではプロラクチンの遺伝子が数個重複している。

図6aは同義置換と非同義置換を各遺伝子の間で計算した結果である。同義置換を横軸に、非同義置換を縦軸にとった。●はヒト-マウスなどの生物種の間で生長ホルモンの遺伝子を比較、□は重複した遺伝子間の比較である。直線は回帰直線である。図から明らかなように回帰係数は重複した遺伝子の方がずっと高い。すなわち同義置換に比べてアミノ酸置換が高まっていることがわかる。図6bはプロラクチンの遺伝子置換に比べてアミノ酸置換が高まっていることがわかる。図6bはプロラクチンの遺伝子で同じ分析をした結果である。やはり重複した遺伝子の比較ではアミノ酸置換が速まっていることが分かる。

このような状況では先程のアミノ酸置換についての集団サイズの影響が異なってくることに注意してほしい。アミノ酸置換は集団の大きさのみでなくむしろ積極的な自然淘汰の働き方に依存することになる。アミノ酸各々の置換については淘汰は弱くほぼ中立の範囲に入るものが多いのではなからうか。遺伝子はDNAの柔軟性に依存しつつ、自然淘汰と遺伝的浮動との相互作用のもとに進化するといえる。

文献

1. E.Zuckerandl and L. Pauling : Evolutionary divergence and convergence in proteins. In" Evolving Genes and Proteins" ,ed. V. Bryson and H.J.Vogel, pp.97 - 166, Aca-

demic Press, New York(1965)

2. M.Kimura : Diffusion models in population genetics. Journal of Applied Probability 1 : 177 - 232(1964) .
3. M.Kimura" The Neutral Theory of Molecular Evolution" ,Cambridge Univ. Press, Cambridge(1983).
4. T.Ohta : The nearly neutral theory of molecularevolution. Ann. Rev. Ecol. Syst. 23, 263 - 286 (1992) .

この論文の一部は「日本の科学者」28巻12号に発表した。