

タンパク質の適応地形の探索

●坂本 健作

東京大学大学院理学系研究科（現；理化学研究所ゲノム科学総合研究センター）

＜研究の目的と進め方＞

個々のタンパク質を、配列空間における適応地形上の個々のピークと見なすとすると、タンパク質の進化は配列空間における「解」の探索と見なすことができる。コンピュータ科学の分野では、「遺伝アルゴリズム (GA)」における「交叉」、「突然変異」、「世代交代」などの操作が探索プロセスをどのように加速するのが明らかにされてきた。本研究では、これらの知見をタンパク質の人工進化実験に応用することの有効性の検証を目的とした。

本研究の特徴は、タンパク質進化のためのGAの手順が分子生物学的操作（ウェットな実験）によって実行されることにある。このようなGAを「ウェットGA」と名づける。GAは本来in silicoに用いられるものなので、最初に、GAにおける概念とタンパク質進化における概念との対応付けをする必要がある。さらに、コンピュータ上の操作と実験操作との対応を見出すことが必要である。コンピュータ上では探索に有利な操作であっても、分子生物学実験としては実施できない操作がある。反対に、タンパク質の進化学において通常用いられる操作でも、そのままの形ではGAの中に対応する操作が存在しないケースもある。

本研究では、情報科学の手法と分子生物学の方法とをつなぐために、「DNA計算」のコンセプトを導入した。DNA計算とは、「DNA上の塩基配列として表現されている情報を、分子生物学の一連の実験によって処理することで、最後に、DNA配列の形で計算結果を得る」というものである。タンパク質の構造・機能の情報は本来DNA上にコード化されているものなので、GAのアルゴリズムを一連のウェットな実験によって実行し、あるタンパク質遺伝子の塩基配列の情報を処理することで、新規な性質を持つタンパク質遺伝子の配列が得られると期待される。

＜研究開始時の研究計画＞

(1)ウェットGAのアルゴリズムの考案： ウェットGAにおいては、コード（DNA配列）と表現型（タンパク質の性質）との対応づけは自然になされ、DNA分子を生物の中で発現させれば容易にその表現型が得られる。DNAを使ってどのようにGAの計算アルゴリズムを実装するかについては様々な考え方ができる。コンピュータ上の操作と実験操作との異同を検討し、GAの探索機能を損なわずに、かつ、分子生物学実験によって実行できるアルゴリズムを考案する。

(2)ウェットGAの有効性の検証： 考案されたウェットGAの有効性を次の2通りで検証する。1つは、コンピュータ実験である。コンピュータでウェットGAをシミュレーションすることは可能であるので、タンパク質とは関係なく、一般的な解探索法としてウェットGAが有効に働くかどうかをコンピュータ実験によって検証する。次に、実際に分子生物学実験によってウェットGAを実行し、タンパク質の人工進化に有効かを検証する。

＜研究期間の成果＞

(1)ウェットGAのアルゴリズムを、「熱力学的GA(喜多ら)」に基づいて考案した。全くランダムなアミノ酸配列から探索をスタートすることは現実的ではなく、通常、野生型タンパク質（＝既存の1つの解）から出発して探索領域を徐々に広げるという方法をとる。よって、ウェットGAでは配列多様性の適切なコントロールが必要である。ここで、「温度」パラメータを、変異体間の活性（適応度）の差異の評価の調整に関わるパラメータとみなせる。

(2)コンピュータ実験：考案したウェットGAのアルゴリズムを用いたコンピュータ実験を行い、あるモデル問題について最適解の探索を行った。パラメータの値を適切に設定すれば、高い確率で最適解を見出すことができた。この探索における配列多様性の推移を調べ、適切に多様性がコントロールされていることを示した。

一点交叉法の開発：コンピュータ実験によって、DNAシャフリングンのような多点交叉よりも、一点交叉の方が探索性能が良いことが示されたので、一点交叉の実験法を開発した。少数点での交叉は、PCRで遺伝子集団を増幅するときに自然と起こる現象なので、交叉した遺伝子のみを選択することがポイントであった。実際に遺伝子間の交叉を行い、交叉位置を解析し、45個の交叉産物の塩基配列を解析し、交叉位置が遺伝子の全長にほぼランダムに分布すること、および、突然変異率が0.1%程度の低率であることを示した。突然変異率が低いので、突然変異と交叉を独立に行うことも可能になった。

＜国内外での成果の位置づけ＞

DNA計算の応用として有望であると評価する研究者もいる一方、通常のタンパク質の進化学と違いが無いとする意見もあった。実際にタンパク質進化へ応用した成果が得られていないので分子生物学の分野では全く評価されていない。

＜達成できなかったこと、予想外の困難、その理由＞

ウェットGAが、タンパク質の人工進化に有効かの検証までは行うことができなかった。時間不足が理由である。

＜今後の課題＞

In silico GAをウェットな実験で完全になぞることではできない。それでもタンパク質の人工進化には有効であるかが問題である。有効である場合、それがタンパク質のn成り立ちに根ざした理由をもつならば、タンパク質の構築原理や進化のメカニズムについて重要な知見が得られる可能性がある。

＜研究期間の全成果公表リスト＞

Sakamoto, K., Yamamura, M., Someya, H. "Toward "wet" implementation of genetic algorithm for protein engineering", Lecture Note in Computer Science 3384, 308-318 (2005).