

Genetic Fusion

池上 高志 (京大基研)

金子 邦彦 (東大駒場)

◆ はじめに

生物種の進化の動的過程は、一般に(突然変異)+(自然選択)をベースにしてモデル化される。今までは主に確率微分方程式による解析が行なわれてきた。しかし分子生物学の急速な発展に伴い、単純な(突然変異)+(自然選択)の図式では済まないことが明らかになってきた。遺伝子レベルのダイナミクスはもっと複雑で、細かくプログラミングされているようである。それらのダイナミクスのうち、『遺伝子融合』というプロセス(e.g. Genetic Fusion)にここでは着目してみた。

例えばある種の遺伝子配列をずっと遡って調べてみた場合、それがひとつの種の遺伝子配列から単に突然変異によって変化してきたわけではない事がわかる。他の種の遺伝子の痕跡、同じDNA配列の無数のコピー、動植物にまたがるヘモグロビンDNAの相同性etc.は同・異種間の遺伝子レベルでの相互作用を示唆している。種および個を規定する遺伝子群の配列パターンは、けっして静的なものではなく、場合によっては種を飛び越えて移動し変化しうるものなのである。

この遺伝子配列の「移動・合体」が Genetic Fusion (GF) である。我々はここでGFをシミュレートする簡単なモデルを提唱し、GFの効用はなにか、GFがどうしても自然界で適用されるか、を探ることにしたい。

■ モデル

ここではDNA配列を簡単に1次元の $[0,1]$ ビット配列で表わすことにする。これらのビット列に対し、環境への適応度関数 W を設定する。適応度とは次世代にどのくらい子孫を残せるか、ということである。種 i の適応度関数を W_i とおこう。この時GFは次のプロセスとして定義される。

ふたつの種 i と j をもってきて、合体させる。この時できた合体種 k の適応度関数 W_k が $W_i + W_j$ より大きければ合体が成立する。

適応度関数は、ビット配列の関数として前もって与えておく。具体的にはSKスピンガラスのエネルギー関数あるいはその変形をもちいるとする。シミュレーションは、初期に適当なビットパターンをもつ小さな断片を用意し、これにMutation + GFを適用する。Mutation (M)とは、配列をランダムに1ビット変化させ、その時配列の適応度が増すならばその変化を許容する、というものである。

★ 結果と考察

モデルのシミュレーションの結果を、各種の適応度の時間発展の様子として描くことができる。斜めに走る太い線は、GF過程を示し、点線はM過程を示す（図参照）。

1. 『モジュール』型の種の存在。

種の時間発展をみていると、それ自体はM過程の作用を受けず、もっぱら他の種のGFパートナーとして活躍している種のいることに気がつく。これらを『モジュール』種とよぶことにする。この『モジュール』種は短いビット長をもち、それ自体のWは大きくないが、GFを通して相手のWを下げるのに寄与する。『モジュール』種が、それ自身と合体してビット長を増していく過程もよくみられる。このような重複によって種が進化していく過程は、自然界にもみられる（大野氏の説）と思われる。例えば自然界のDNA配列に何故同じパターンが繰り返し存在しているのか、といったことの説明になるかもしれない。

2. 断続平衡。

グールドらが、指摘したような断続平衡過程がわれわれの結果にもみられる。単にM過程だけだと段々に進む進化（より大きい適応度の獲得）が、GFをいれることで、断続的な飛びが出現する。これは、1bit mutationにより、自分の近傍で適応度のよりよい配列を探すが、いいパートナーをみつけたら（例えば先の例にあげた『モジュール』種）そいつと合体して一気に適応度をあげる、というGFプロセスの結果である。

3. 実際の効用

GFはさらに実際の効用をもっている。つまり適応度関数に対する最適化問題を解く高速アルゴリズムを与えていると考えられる。多くの準安定状態をもつ適応度関数の場合、単なるM過程だと深い谷間にはまれば出られなくなる。このため多数の初期パターンを用意して探さねばならないが、ものによっては計算時間が発散してしまう。これに対しGFの場合、深い谷間からの脱出が『モジュール』種の存在により可能となるわけだ。もっともGFで構成されるパターンは繰り返しパターンとなりやすい。したがって最適パターンが、部分的繰り返しパターンによってよく近似される場合は、GFが特に効果的といえよう。

今後このモデルを拡張して、部分配列を切り離すプロセスを取りいれたり、遺伝子レベルの共生のモデルを考えていく予定である。

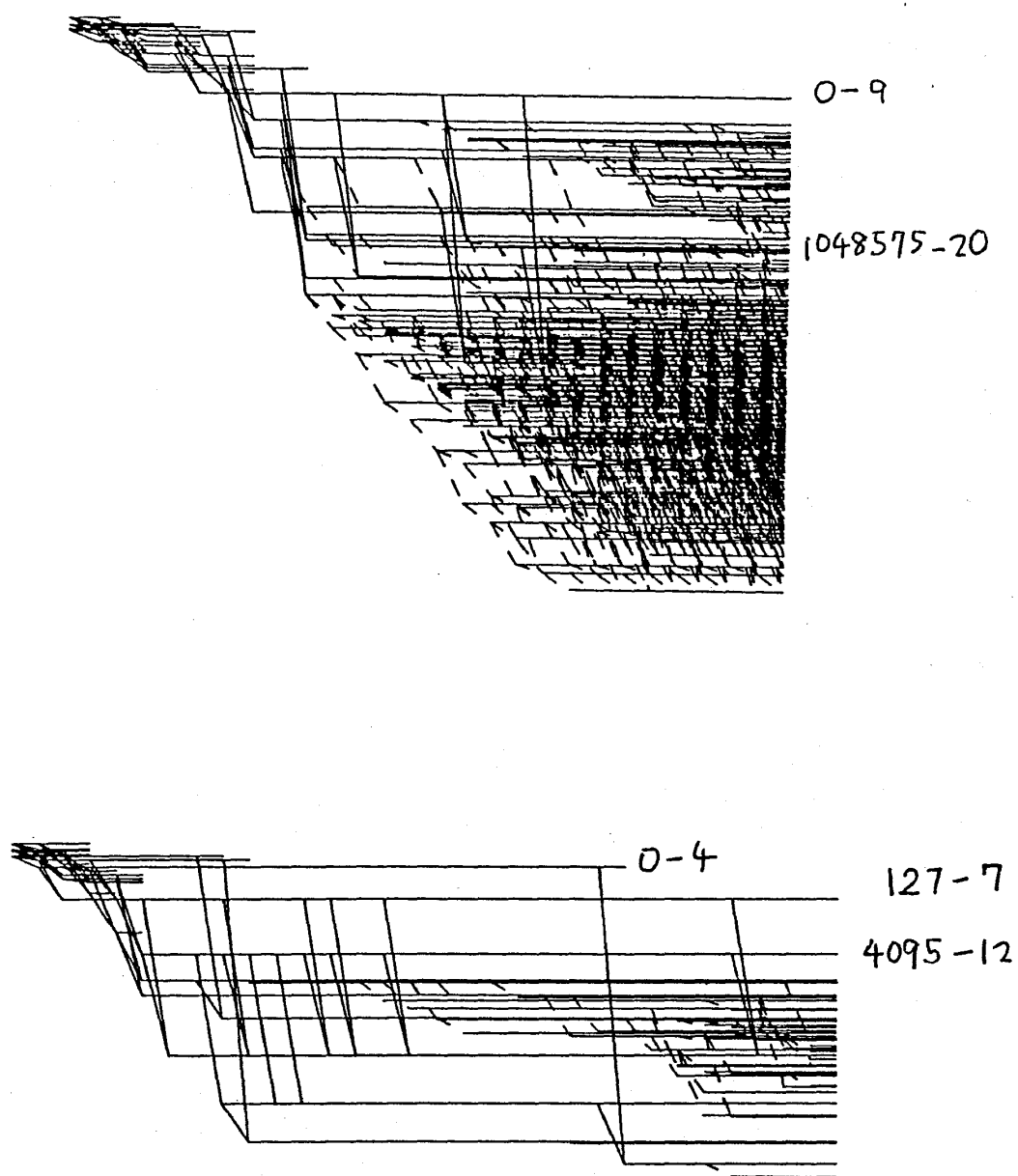


図 右方向に時間。 下向きに適応度が増す。 斜め下向きの太い線が Genetic Fusionを示し、点線はMutation。 横についている数字は、種のコード名。 例えば上図では種<0-9>が、下図では種<127-7>が『モジュール』種として振る舞っている。