

原核生物の遺伝子水平移動と可動性遺伝因子

元来、交配が可能かどうかで定義された種という概念は、当然のことながら有性生殖を行わない生物には適用できない。すなわち、原核生物における「種」の意味するところは我々哺乳類の「種」の意味とは全く異なる。しかしそれ以上に、同一の遺伝子セットが時間とともに変化することで分岐する我々の種と異なり、原核生物では、持っている遺伝子セットそのものが大きく変化することがわかってきた。真正細菌と古細菌では、遺伝子水平転移（水平伝播）が頻繁に起こっているため、持っている遺伝子セットは近縁な種、株であっても大きく異なる場合がある。例えば、配列の決定された3株の大腸菌 (*Escherichia coli* strains CFT073, EDL933, MG1655) 全てに共通する遺伝子は全体のたった40%しかない。また、多くの薬剤耐性菌の耐性獲得は耐性遺伝子をコードしたプラスミドの獲得やそのプラスミドのゲノムへの組み込みによって起こる。更に、原核生物には環境中のDNAを積極的に取り込む機構が存在しており、突然変異のDNA源として利用されている。

遺伝子水平転移の検出には様々な手法が提案されているが、大きく3つに分類できる。まず、特定の遺伝子の系統樹を作成し、生物が大きく異なっているのに対して遺伝子同士が近縁な場合は水平転移によって遠縁の生物に移った可能性が高い。また、非常に近縁な2種類の生物を比較すれば、どこに挿入が入り、どこに欠失が起こったかがわかるはずである。系統の離れた原核生物のゲノム配列が大量に得られる今日では、ゲノムの配列上の特徴を利用して水平転移遺伝子を検出する手法が最も一般的となっている。ゲノム配列はただの文字情報ではなく、DNAという化合物として物理的な性質を持っており、生物の環境に応じて様々な影響を受ける。例えば、温度の高い環境を好む好熱菌では、塩基対の結合力が強いGC塩基対がAT塩基対に対して安定なため増加する傾向がある。また、同じアミノ酸をコードしている縮重コドンでもtRNAの量の違いにより生物ごとにコドンの偏りが生まれる。これらの特徴を利用して、最近水平転移してきた「余所者」の遺伝子を検出するのである。簡単なものでは、ゲノム全体の平均的なGC含量に比べて明らかにGC含量が多い／少ない遺伝子は水平転移によって最近移ってきた可能性が高い。より複雑には、コドンの各位置についてGC含量を調べたり、ジヌクレオチドやトリヌクレオチド（コドン）の使用頻度を見たりする手法も示されている。Garcia-Vallveらはこれらの解析から予測される水平転移遺伝子のデータベースを作成している (Garcia-Vallve et al. 2003、<http://www.fut.es/~debb/HGT>)。が、論文掲載後はあまり更新されていない。

Nakamuraらは、ペンタヌクレオチド（5塩基）の頻度を利用してどのような遺伝子が水平転移しているのかを網羅的に調べた (Nakamura et al. 2004)。彼ら

の解析では、116 種の原核生物ゲノムの、平均して 14%程度の遺伝子が最近水平転移をしたと結論付けている。多いものでは全遺伝子の 1/4 が最近水平転移してきた生物も存在する。彼らは、まず、ゲノム配列を蛋白質をコードしている領域とコードしていない領域に分け、連続する 5 塩基の頻度を計算した。そこからベイズの公式を用いて、与えられた配列が特定の生物のゲノムの蛋白質をコードしている領域である確率を計算した。蛋白質をコードしている遺伝子においてこの確率が低いとその遺伝子は水平伝播された確率が高い。そして、この確率が高いとその遺伝子が生物本来の遺伝子である確率が高い。このため、この手法の一つの特徴として、もし、配列の読まれた生物の中に、水平転移した遺伝子の由来する生物やその近縁種が含まれていた場合には、その由来が推測可能である。しかし、彼らの研究は、むしろ水平転移した遺伝子の種類を詳しく解析したところに特徴がある。

Nakamura らは水平転移した遺伝子を TIGR の定義している機能カテゴリごとに分類した。最も頻度が高かったのはプラスミド、ファージ、トランスポゾンに関連する機能 (Plasmid, phage, and transposon functions) を持つ遺伝子だったが、これらは水平転移しやすい遺伝子である (後述) のでその後の解析からは除外している。次に多かったのが細胞膜関係 (Cell envelope)、そして、調節機能 (Regulatory functions)、細胞プロセス (Cellular processes) と続く。更に細かく見ると、細胞膜関係では、表面構造蛋白質 (Surface structures)、多糖やリポ多糖の生合成と分解 (Biosynthesis and degradation of surface polysaccharides and lipopolysaccharides) が多く、細胞プロセスでは抗生物質合成などの疾病関係遺伝子 (Pathogenesis) が多い。これらの遺伝子群は感染性や疾病に関連している。疾病関連遺伝子はまとまって転移し、Pathogenicity island と呼ばれるクラスターを作ることが知られており、これまでの知見と合致している。また、生物の遺伝子は水平転移しやすい作用遺伝子群 (Operational genes) と水平転移しにくい情報遺伝子群 (Informational genes、転写翻訳など) とに分けられることがこれまでも知られていたが、彼らの解析から、作用遺伝子群は、水平転移しやすい DNA 結合や疾病関連遺伝子などと、水平転移しにくいアミノ酸、補酵素、塩基などの合成やエネルギー生成などに分けられることが新たにわかった。彼らもまた、水平転移した遺伝子をまとめて公開している (http://poplar.genes.nig.ac.jp/~hgt/viewer_top.cgi)。

多くの水平転移にはプラスミド、ファージ、トランスポゾンなどの可動性遺伝因子が関わっていると考えられている。実際、Nakamura らの解析でも水平転移の頻度が最も高いのはこれらの因子のコードする遺伝子である。Frost らの総説では、これらの因子と水平転移との関わりをまとめている (Frost et al. 2005)。

プラスミドが遺伝子の細胞間移動に関わっていることは古くから知られてい

る。大腸菌のプラスミド F 因子は性プラスミドとして知られ、F 因子を持つ細胞から繊毛が伸び、F 因子を持たない細胞へと自身の DNA を送り込む。F 因子はたまた大腸菌のゲノム DNA へ挿入されることがある。ゲノムに挿入された F 因子を持つ細胞から F 因子を持たない細胞へ DNA が侵入する際には、F 因子のみならず宿主の DNA も一緒に移行するため、遺伝子の細胞間移動が起こる。プラスミドの細胞間移動には、Conjugative（接合）遺伝子、あるいは Transfer（トランスファー）遺伝子と呼ばれる遺伝子群が関与する。代表的な接合遺伝子には、接合相手を決めて DNA を送り込む Coupling protein や DNA にニックを入れる Relaxase がある。Coupling protein は TraG-like family に属する ATPase で細胞膜に局在する。Relaxase は *nic* site と呼ばれる特定の配列を認識し一本鎖切断する。グラム陰性菌の接合では、繊毛も重要であり、この繊毛を通じて一本鎖 DNA が受け手の細胞へ移行する。繊毛の形成と DNA の移行にはプラスミド自身がコードするタイプ IV 分泌系が働いている。

バクテリオファージ（ファージ）は、一本鎖あるいは二本鎖の DNA または RNA と様々な形態で存在し、大きさも数キロから数百キロと幅広い。ファージには宿主の DNA 複製系を盗用するための特別な遺伝子群と、DNA を包み込むキャプシド蛋白質が必須である。ファージはキャプシド蛋白質に包まれた DNA の形で別の細胞に感染することができる。ファージは宿主内で独立した複製ユニットとして増殖する場合と、宿主の DNA に組み込まれてプロファージとして存在している場合がある。この組み込みは、特定の配列の間へファージ DNA 全体が挿入されることによって起こり、ファージのコードする Tyrosine recombinase あるいは Serine recombinase が触媒する。IS などのトランスポゾン自身は他の細胞に転移することはできないが、コードする DDE transposase によって宿主ゲノムに挿入される。

原核生物の遺伝子水平転移には、プラスミドやファージによる DNA の細胞間移動と、ファージやトランスポゾンのゲノムへの挿入機構の 2つのステップが存在すると考えてよいだろう。巨大なプラスミドにはトランスポゾンやファージが挿入されていることもあり、プラスミドの細胞間移動後、挿入されていたトランスポゾンが転移してゲノム中に挿入されるなど様々な遺伝子水平転移のステップが想定されている。しかし、それ以上に重要なことに、遺伝子の水平転移が起こるためには、まず可動性遺伝因子の転移単位あるいは複製単位の内部に遺伝子が挿入されていなければならない。このステップに関しては知見が非常に少ないのが事実のようだ。唯一確かな機構は、同じ種類の転移因子が近傍に挿入されていた場合に、2つの転移因子が挟まれている遺伝子ごと転移するという機構で、抗生物質耐性遺伝子がこのような機構によって水平転移している。

Frost らは現在の可動性遺伝因子解析の課題、問題点として、機能未知の遺伝子が多いこと、遺伝子の命名基準が無いことなどを挙げている。従って機能未知の遺伝子にはかなりの数の可動性遺伝因子に関わる遺伝子が含まれていると考えられる。ところで、多くの可動性遺伝因子には、遺伝因子の可動性に関わらない機能を持つ遺伝子が多く含まれている。この内、どこまでが可動性遺伝因子の遺伝子で、どこまでがその中に偶然挿入された宿主の遺伝子であるのかを区別するのは不可能である。実態として可動性遺伝因子と宿主の遺伝子との間には相互に移入があり、その動的な構造が遺伝子の水平転移を可能にしている。そしてこれが生物同士の遺伝子の移入を生んでいるのである。

Garcia-Vallve S, Guzman E, Montero MA, Romeu A.

HGT-DB: a database of putative horizontally transferred genes in prokaryotic complete genomes.

Nucleic Acids Res. 2003 Jan 1;31(1):187-9.

Nakamura Y, Itoh T, Matsuda H, Gojobori T.

Biased biological functions of horizontally transferred genes in prokaryotic genomes.

Nat Genet. 2004 Jul;36(7):760-6.

Erratum in: Nat Genet. 2004 Oct;36(10):1126.

Frost LS, Leplae R, Summers AO, Toussaint A.

Mobile genetic elements: the agents of open source evolution.

Nat Rev Microbiol. 2005 Sep;3(9):722-32. Review.

2005/12/27

小島 健司 著
禁 無断複写転載