

そして誰もいなくなった：ヒトゲノムの中の DNA トランスポゾン

LINE はヒトゲノムの 21%を占める。続いて大きなグループの反復配列は SINE で 13%を占める。レトロウイルス由来の転移因子である内在性レトロウイルス (endogenous retrovirus) は 8%を占める。これらは全て逆転写によって RNA から DNA を作って増幅するレトロトランスポゾンである。それらに比べると足してもたったゲノムの 3%にしかない DNA トランスポゾンが軽視されてきたのはある意味当然であろう。また、DNA トランスポゾンの転移に由来する遺伝病が見つかっていないのも、ヒトの DNA トランスポゾンの研究が進んでいない理由である。しかし、ヒト以外の生物を含めて考えると、DNA トランスポゾンの研究はレトロトランスポゾンの研究よりもはるかに進んでいるのも事実である。

Pace らはヒトゲノム中の DNA トランスポゾンの網羅的な解析を 3つの手法を併用して試みた (Pace and Feschotte 2007)。まず、RepeatMasker でヒトゲノムを解析した結果から、DNA トランスポゾンだけを抽出した。ヒトゲノムに見つかった DNA トランスポゾンは 4つのスーパーファミリーに分類された。すなわち、hAT、MuDR、*piggyBac*、そして *Tc1/mariner* である。hAT は *hobo*、Ac、Tam3 を含み、それらの頭文字から命名されたグループで、挿入の際に両側にできる直列反復配列、すなわち target site duplication(TSD)が 8塩基という特徴がある。また、トランスポゾンの両末端に見られる逆位反復配列 terminal inverted repeat(TIR)は短い。MuDR は 9塩基の TSD を作り、長い TIR を持つ。*piggyBac* の TSD は TTAA の 4塩基である。*Tc1/mariner* はその名の通り *Tc1* と *mariner* を代表とするグループで、TSD は TA の 2塩基である。TSD の長さと TIR は新規の DNA トランスポゾンがどのスーパーファミリーに属すかの指標となる。ヒトゲノムでは、hAT が全 DNA トランスポゾンの 2/3 程度、*Tc1/mariner* が残り 1/3 程度を占め、残り 2つのスーパーファミリーのトランスポゾンは 1%にも満たない。

Pace らはヒトゲノム中の DNA トランスポゾンの転移時期を調べるために、レトロトランスポゾンの L1 や *Alu* に挿入されているものを探索した。L1 や *Alu* ではこれまでに詳しい転移時期の解析がなされており、それらに挿入されていることでそれ以後に転移した事を示すことができる。霊長類だけに見られる L1 のサブファミリーに挿入されていた DNA トランスポゾンは 10種類 19挿入見つかった。L1 のコンセンサス配列は知られているのでそこからトランスポゾンの両側にできる TSD を見つける事ができる。実際、9種類では TSD が見付き、それぞれのスーパーファミリーの特徴と一致した。DNA トランスポゾンに挿入されている L1 で最も新しいサブファミリーは L1PA8A で約 4,000 万年前に転移し

ていた。同様に *Alu* に挿入されていた DNA トランスポゾンが 7 種類 15 挿入、内在性レトロウイルスに挿入されていたものが 6 種類 18 挿入見つかった。共通するものが多いのでまとめると 11 種類 (MER1A、1B、30、30B、75、75B、85、107、MADE1、HSMAR1、Charlie3) の DNA トランスポゾンが霊長類のレトロトランスポゾン内に挿入されていた。

続いて ENCODE 計画に含まれているヒトゲノム配列中に挿入されている 794 の DNA トランスポゾンが他の哺乳類ゲノム配列中の相同な位置にあるかどうかを解析した。これらはヒトゲノム中に見つかる 125 の DNA トランスポゾンファミリーの内、111 種を含んでいる。その内、10 種 (MER1A、1B、30、30B、75、85、107、MADE1、HSMAR1、Charlie3) では、ヒト、アカゲザル、マーモセットで挿入が見られるが、オオガラゴでは挿入されていない座位が見つかった。DNA トランスポゾンは上述したように挿入の際に短い配列が重複し、TSD となる。逆に切り出しの際には TSD の片方と自身の配列を切り出す。しかし正確に切り出される確率は低く、周囲の配列を含めて切り出すことが多い。従って挿入の有無を比較した時に DNA トランスポゾンの両側に TSD がきれいに見つければ、片方の系統で挿入されたことが推察できる。比較の結果、オオガラゴで抜けたわけではなく、ヒト、アカゲザル、マーモセットの系統で挿入されたことが推察できた。ここまでの結果は、レトロトランスポゾンに挿入されていた DNA トランスポゾンの解析結果と一致する。前の解析では霊長類で挿入されていることが確認できたもう 1 種類の DNA トランスポゾンである MER75B については、ガラゴで相同な位置が欠失しており、挿入の有無を確かめられなかった。

上述の 11 種類を除くと、霊長類以外では挿入されていない DNA トランスポゾンは 23 種類 163 挿入みつかった。これらはガラゴで挿入されている座位が見つかるので、霊長類の共通祖先で転移したものが含まれている。他 77 種類の DNA トランスポゾンは霊長類以外でも相同な位置に挿入が見つかった。ENCODE に含まれていなかった 14 種類の DNA トランスポゾンは個別に同様の解析を行ない、内 6 種類は霊長類でだけ挿入が見られた。

種間のゲノム比較の結果からは、40 種の DNA トランスポゾンが霊長類が生まれてからも転移していた事がわかった。これらの 40 種の DNA トランスポゾンはヒトゲノム中に 98,300 コピー存在する。ただし、それ以前から転移活性を持っていた事も考えられるので、これら全てが霊長類で転移したとは限らない。

彼らが用いた 3 つ目の手法は、これまでの 2 つとは異なり、挿入の有無を利用したものではない。コンセンサス配列からの塩基置換の量を基にして転移年代を推定する手法である。すなわち同じ種の DNA トランスポゾンの多様性が低いものほど最近転移活性があったことを示している。最も変異の蓄積が少な

ったのは MER85 で平均転移年代は 3,700 万年前と推定された。この年代は、旧世界ザルと新世界ザルとが分岐した時期とほぼ一致する。

さて、ここまでの 3 手法の結果をまとめよう。最初の手法で霊長類特異的な挿入が見つかった DNA トランスポゾン、MER75B を除いては、次の手法で直鼻猿類（類人猿、旧世界猿、新世界猿をまとめたグループ）で転移した DNA トランスポゾンと共通していた。すなわち、MER1A、1B、30、30B、75、85、107、MADE1、HSMAR1、Charlie3 の 10 種である。2 番目の手法では見つからなかったが、MER75B も直鼻猿類で転移していたと考えられる。この内自前で転移酵素（transposase）をコードするものは、HSMAR1 と Charlie3 の 2 つのみで、他はコードしない。従って前 8 種の転移には別の DNA トランスポゾンから転移酵素が供給されなければならない。MADE1 は HSMAR1 と同じ Tc1/*mariner* スーパーファミリーに属す。実際、MADE1 は HSMAR1 の内部欠失によって生じた非自律型トランスポゾンである（RepBase 参照）。MER1A、1B、30、30B、107 は Charlie3 と同じ hAT スーパーファミリーに属す。MER1A と MER1B は Charlie3 の内部欠失によって生じた非自律型トランスポゾンであり、Charlie3 が転移させていたと考えるのが自然である。MER30、30B、MER107 は Charlie12 に由来する非自律型トランスポゾンであるが、Charlie12 がヒトゲノム中に 2 コピーしかないためか、解析に含まれておらず、Charlie12 が MER30、30B、107 を最近転移させていたのかどうかは不明である。

残る 3 つ MER75、75B、85 は *piggyBac* スーパーファミリーに属す。*piggyBac* スーパーファミリーに属する転移酵素を持つトランスポゾンでヒトゲノム中に見つかるものは Looper のみでその転移年代はこの解析では 7,700 万年前と推定された。転移酵素をコードしていても TIR が変異を起こして転移酵素が認識できなくなってしまうとその DNA トランスポゾンは転移する事ができなくなる。もう転移できない DNA トランスポゾンが転移酵素を提供し、他のトランスポゾンを転移させることは一般に見られる。従って MER75、75B、85 は転移できなくなった Looper が転移酵素を供給して転移させていたのかもしれない。あるいは Looper に近縁な別の自律性転移因子が存在したのかもしれない。

まとめると、直鼻猿類の共通祖先で転移していた DNA トランスポゾンは以下のようになる。

Superfamily	autonomous	non-autonomous
Tc1/ <i>mariner</i>	HSMAR1	MADE1
hAT	Charlie3	MER1A,1B
	Charlie12?	MER30,30B,107
<i>piggyBac</i>	Looper?	MER75,75B,85

霊長類の共通祖先ではもっとたくさんのDNA トランスポゾンが活性を持っていた。論文では合計 40 種類と描かれているが、数えてみると 39 種類しか見つからない。おそらく Ricksha を二重に数えているか、Ricksha と Ricksha_a、Ricksha_0 を間違えているのであろう。上記のトランスポゾンを除くと 28 種のトランスポゾンが霊長類の進化初期には転移していたはずである。内、MER82 は Tigger2 の、MER46A,B は Tigger4(Zombi)の、MER44A,B,C,D は Tigger7 の、内部が欠失した因子である。MER2,2B も Tigger7 とよく似ている。

Superfamily	autonomous	non-autonomous
Tc1/mariner	HSMAR2,Tigger1,2,2a,3(Gole m),3b,4(Zombi),5,5a,5b,7	MER2,2B,44A,44B,44C,44D,46A,46B,6,6A,6B,6C,8,82
MuDR		Ricksha, Ricksha_b, Ricksha_c

このように霊長類の進化初期には多数のDNA トランスポゾンが転移していた。しかし、Pace らの解析では、結局、直鼻猿類の内部で新たに転移したDNA トランスポゾンは1 つも見つからなかった。また、配列の多様性からも、最も新しいトランスポゾンでも旧世界猿と新世界猿との分岐以前に転移活性は失われてしまった事が示唆された。おそらくこの時期に最後の自律性のDNA トランスポゾンの転移酵素が破壊されたのだろう。そして直鼻猿類のゲノムからはDNA トランスポゾンは誰もいなくなった。

Pace JK 2nd, Feschotte C.

The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage.

Genome Res. 2007 Apr;17(4):422-32.

2009/03/12

小島 健司 著

禁 無断複写転載