

Non-LTR レトロトランスポゾンの翻訳戦略：1 本の RNA から 2 種類の蛋白質

教科書にも書かれているとおり、真核生物の mRNA は 1 種類の蛋白質だけをコードしている。これをモノシストロニック (monocistronic) RNA と呼ぶ。これに対して原核生物の RNA は複数の蛋白質をコードしており、このような RNA をポリシストロニック (polycistronic) RNA と呼ぶ。特に 2 種類の蛋白質をコードしている場合をバイシストロニック (bicistronic) RNA、あるいはジ (ダイ) シストロニック (dicistronic) RNA と呼ぶ。この違いは、主に翻訳開始の機構が異なることに起因している。すなわち、原核生物では、シャイン・ダルガーノ (SD) 配列をリボソームが認識して直後の開始コドンから翻訳が開始される。SD 配列は RNA 上のどこにあっても良い。一方、真核生物では、mRNA の 5' 末端のキャップ構造を認識し、そこから mRNA 上を滑っていった最初の開始コドン AUG から翻訳を開始する。真核生物では実際にはさまざまな転写開始因子が複雑に絡み合ってこのような認識を行なっている。この違いから、真核生物の mRNA は、たとえ複数の蛋白質をコードしていたとしても、最も 5' 側にコードされている蛋白質しか翻訳されない。

Non-LTR レトロトランスポゾンの RNA は蛋白質をコードする mRNA であると同時に、逆転写の鋳型となるゲノム RNA でもある。これは一見便利のように見て大きな制約でもある。RNA には全ての遺伝子が乗っていなければゲノム RNA として使えない。一方で、RNA 上には一つの蛋白質しかコードできない。この、mRNA 兼ゲノム RNA にかかる制約は RNA ウイルスでも当てはまる。非常に単純な手法だが、複数の蛋白質を一続きで翻訳し、その後切断してそれぞれの蛋白質を取り出す、という手法を多くの RNA ウイルスは採用している。この場合は厳密には複数の蛋白質を翻訳しているわけではないが、実際に複数の蛋白質を翻訳する機構もいくつか知られている。

まず、これまでに知られている真核生物のポリシストロニック RNA 翻訳機構を列記しよう。参考文献は省略する。

(1) Leaky scanning

開始コドンとして認識されにくい文脈中の AUG を読み飛ばして、下流のより認識されやすい AUG から翻訳を開始する。

(2) Ribosome shunt

リボソームがスキャン中に、RNA 二次構造を跳び越えることで二次構造中の AUG を無視する。

(3) Translation termination reinitiation

非常に短い ORF (1-3 コドン) を翻訳したリボソームが下流の AUG から翻訳を再開する。

(4) Internal ribosome entry

キャップ非依存的に、特殊な RNA 二次構造 (IRES) を認識してリボソームが結合し、翻訳を開始する。

(5) Translational coupling

上流の ORF の翻訳と下流の ORF の翻訳が強く連関している。上流の翻訳が下流を抑制する場合と、促進する場合の両者ともが Translational coupling と呼ばれるが、ここでは、後者のみを示すのに用いる。この場合、上流の ORF を翻訳したりリボソームだけが下流の AUG から翻訳を再開する

(6) Ribosomal frameshifting

上流の ORF の翻訳途中にリボソームがフレームを間違え、下流の ORF まで続けて翻訳する。

(7) Readthrough

終止コドンに稀に suppressor tRNA が認識することでその下流まで翻訳する。

1-3 の機構では、複数の蛋白質は合成されない。1 つの大きな蛋白質と少数の短いペプチドが合成されるだけである。1 は、N 末の構造を少し変えた (例えばシグナルペプチドの有無) 複数の isoform を合成するのに用いられる場合もある。4-7 が複数の蛋白質を合成する機構である。どれもウイルスが使っている機構である。4 はピコルナ様ウイルスの Cricket Paralysis Virus、5 はパラミクソウイルスなどいくつかの RNA ウイルス、6-7 は HIV などのレトロウイルスが利用している。ちなみにほとんどの Internal ribosome entry site/sequence (IRES) は 5'UTR にあり、ポリシストロニック RNA の翻訳には関与していない。更に、ポリシストロニック RNA と間違えやすい機構として、

(8) Cryptic promoter

弱い転写活性のあるプロモーターが主な転写産物の中ほどにあり、下流だけからなる RNA が少量転写される。

が挙げられる。

話を戻そう。HIV や HTLV などのレトロウイルスでは、Pol 蛋白質の翻訳に 6 の Ribosomal frameshifting が用いられていることが報告されている。これは、同じ塩基の繰り返しによりリボソームが滑りやすくなった配列と、下流の RNA 二次構造の立体障害によって引き起こされる。一方、酵母の LTR レトロトランスポゾン Ty1 と Ty3 では、稀な tRNA を用いるコドン上でリボソームが待ちきれずにスリップすることでフレームシフトが起こる。また、レトロウイルスには Readthrough を使って翻訳を行なうものも知られている。

Non-LTR レトロトランスポゾンの発見はレトロウイルスや LTR レトロトランスポゾンよりも遅く、機構の解析も後を追う形で進められてきた。non-LTR レトロトランスポゾンの ORF1 蛋白質はレトロウイルスや LTR レトロトランスポ

ゾンの Gag 蛋白質に、ORF2 蛋白質は Pol 蛋白質に対応する。このため、実際に解析が行なわれるまでは、non-LTR レトロトランスポソンのバイシストロニック RNA 翻訳機構も LTR レトロトランスポソンやレトロウイルスと同様の機構が想定されていた。しかし、実は全く違う機構を使っていることが Kojima (注：筆者)らの解析から明らかとなった (Kojima and Fujiwara 2005)。Kojima らが解析に用いたのは、カイコのテロメア反復配列にだけ挿入される non-LTR レトロトランスポソン、SART1 である。ヒトの L1 などと異なり、SART1 ではそのほとんどが全長のコピーであることが知られていた。また、転写が SART1 の最初の塩基から正確に始まり、全長の RNA だけが合成されることも明らかになっていた。つまり、5'側が欠失したコピーから転写される RNA や内部のプロモーターから転写される RNA のような、5'末端に近いところに ORF2 の開始コドンがある RNA から ORF2 蛋白質が翻訳されている可能性を除去することができる。SART1 では全長の RNA から確実に 2 種類の蛋白質が翻訳されているはずなのである。

Kojima らは SART1 をカイコと同じ鱗翅目昆虫由来のウイルスである AcNPV に組み込み、同じく鱗翅目由来の培養細胞である Sf9 で大量発現させることにより、SART1 の ORF2 蛋白質の翻訳機構を解析した。これには 2 つの意味がある。一つは、カイコに近縁な系を用いることでより自然に近い解析ができることである。そしてもう一つは、確立された蛋白質発現系であるバキュロウイルス系を応用することでこれまで誰も検出できなかった ORF2 蛋白質を検出することが期待されることである。理想的なバキュロウイルス発現系では、細胞の合成する実に 50%もの蛋白質が目的蛋白質になりうる。実際、彼らはこの条件で ORF2 蛋白質の検出に成功している。ただし、検出できたのは C 末側を途中で切ってヘマグルチニン (HA) タグを取り付けた場合だけで、全長に HA タグを取り付けた場合には検出はできていない。ここで重要なことは、検出した ORF2 蛋白質は明らかに単独で発現していたことである。つまり、ORF1 蛋白質との融合蛋白質として合成されているわけではないことがこの研究により初めて証明されたのである。

続いて Kojima らは、ORF1 と ORF2 の境界領域にさまざまな変異を導入して ORF2 蛋白質の翻訳量の変化を観察した。結果をまとめると、(1) ORF2 の翻訳は AUG から始まる。(2) UAAUG あるいは AUGA のような ORF1 の終止コドンと ORF2 の開始コドンとのオーバーラップが ORF2 の翻訳に必須である。(3) ORF2 の翻訳は ORF1 の翻訳が正常に終結することに依存している。(4) ORF2 の効率的な翻訳には下流の RNA 二次構造が必要である。ORF1 の終止コドンを本来よりも前側に移動させた場合には、ORF2 の翻訳が起こらないという結果から導かれた (3) の事実は、Internal ribosome entry による翻訳開始の可能性を明

確に排除している。同時に弱いプロモーターにより ORF2 単独の RNA が転写されている可能性をも否定できる。更には、ORF1 の翻訳と ORF2 の翻訳とが競合的に起こることを示唆する Ribosomal shunt や Leaky scanning も排除できる。となると残るのは Translation termination reinitiation と Translational coupling であるが、この2つの違いは、前者が ORF1 の終止コドンと ORF2 の開始コドンとの間にある程度の距離が必要なものに対して、後者は距離が開くほど翻訳効率が低下する点にある。終止コドンと開始コドンとがオーバーラップした UAAUG や AUGA も Translational coupling に特徴的な構造である。従って、SART1 の ORF2 翻訳機構は Translational coupling であることがわかる。Translational coupling はもともと原核生物で知られている現象で、真核生物では他には RNA ウイルスでしか知られていない。SART1 が Translational coupling を使用していることは、真核生物のゲノムにコードされている RNA でも Translational coupling によって翻訳されるものがある可能性を強く示している。実際、マウスの Glutamic acid decarboxylase の胚特異的な RNA variant に、2つの蛋白質を1つの RNA から翻訳しているものが知られており、これが UGAUG のオーバーラップした翻訳終止—開始コドンを持っている。これが Translational coupling によって翻訳されている可能性は高い。

同時に Kojima らは、多くの non-LTR レトロトランスポゾンが UAAUG や AUGA を持っていることを示している。一方で、持っていないレトロトランスポゾンが多く存在することもまた事実である。これらは他の機構によって翻訳を行なっていると考えられる。この他の機構を示す論文が Alisch らと Li らによって発表された (Alisch et al. 2006; Li et al. 2006)。

Alisch らは既に確立されていたヒトの L1 の培養細胞転移モデル系を用いて ORF2 の翻訳を解析した (Alisch et al. 2006)。この転移系では、L1 の転移が起これるとネオマイシン耐性遺伝子が発現し、薬剤中の生存細胞数を数えることで転移効率を知ることができる。ORF1 に違いが無い場合には、この転移効率は ORF2 の翻訳効率に等しいと考えてよい。彼らの解析では ORF2 の翻訳はこのように間接的にしか調べられていない。

順を追って結果とそこから導かれた結論を述べよう。ORF1 と ORF2 との間の領域の配列や長さを変化させた場合には、転移効率は低下するが、完全には失われなかった。一方で、ORF1 の終止コドンを変異させると転移効率はほとんど失われた。すなわち、ORF2 の翻訳は ORF1 の翻訳終結の位置に依存する。ORF2 の最初の AUG の後ろに終止コドン进行挿入すると転移はほとんど見られなくなったが、直前に挿入した場合には正常な L1 と同程度の転移が確認された。すなわち、ORF2 の最初の AUG から翻訳は開始される。しかし、この AUG を他のアミノ酸をコードする配列に変化させた場合にもかなりの頻度の転移が見られる。

一方、終止コドンに変化させた場合にはほとんど転移が見られない。すなわち、ORF2 の翻訳は常に同じ位置から始まり、開始コドンが必要としない。続いて、ORF1 と転移する RNA を他から供給して ORF2 の翻訳に対する ORF1 の配列の影響を調べた。ORF2 の上流の ORF1 を EGFP に入れ替えても転移効率はあまり変化しなかった。以上の結果からわかることは、ORF2 の翻訳開始点は ORF1 の翻訳の終了位置と、AUG の周辺の、おそらく下流の塩基配列によって決められており、その際、開始コドンは補助的な役割しか果たさない。他のコドンからもある程度の翻訳開始が起こる。その一方で、ORF2 のモノシストロニック RNA では翻訳は AUG に強く依存する。これは、他のコドンからの翻訳開始は、ORF1 の翻訳終了後の翻訳再開の場合に限られる現象であることを示している。彼らは他にも、ORF2 の上流に強固な RNA 二次構造を挿入した場合に翻訳効率が低下することや、この特殊な翻訳機構が培養細胞の種類に因らないこと、マウスの L1 でもほぼ同様の現象が起こること、本当の AUG に変異を入れ、近くに AUG を新しく作った場合には新しい AUG から翻訳が起こることなどを示している。

これらの結果を踏まえて、Alisch らは L1 の ORF2 の翻訳開始のモデルを提案している。ORF1 の翻訳終了後、リボソームは RNA 上をスキャンし続け、周辺配列の影響のもとで、ORF2 の最初の AUG から翻訳を開始する。この現象は翻訳開始複合体非依存的であると考えられる。また、ORF1 と ORF2 の間の配列の長さを変化させると翻訳効率が変化し、ORF1 の翻訳終結を早めると翻訳がほとんど起こらなくなるという結果からは、スキャンする配列の長さにはある程度の制限があることがわかる。このモデルは、既知の翻訳再開機構とは異なっている。Translation termination reinitiation では、ORF1 が短いのはもちろんだが、AUG から翻訳が再開される。また、Translational coupling でも、AUG から翻訳が開始される。ただし、SART1 では、弱い翻訳開始能のある ACG からだとある程度の翻訳が起こっているので、これは程度の問題に過ぎないのかもしれない。L1 では ORF1 と ORF2 の距離を短くしても翻訳効率は多少ながら低下した。これも、Translational coupling では距離が長くなると翻訳効率が低下することが想定されるのと異なる結果だが、効率の低下が限定的であることを考えると Translational coupling の可能性を完全に排除することはできない。個人的な感想を言うと、SART1 と L1 では、細かな要素の影響の強さが違うだけで類似した機構を使用しているのではないだろうか。

Li らもマウスの L1 を材料とした。彼らはマウスの L1 の翻訳機構は Internal ribosome entry ではないかと当たりをつけて解析を行なっている。ルシフェラーゼアッセイの系を用いて、5'側にウミシイタケのルシフェラーゼ、3'側にホタルのルシフェラーゼを配置し、間に IRES と想定される、ORF1 と ORF2 の開始コドンのそれぞれ上流域を挿入したコンストラクトを作成した。培養細胞での解

析結果では、ORF1 の開始コドンの上流 300nt、ORF2 の開始コドンの上流 200nt に IRES の存在が確認された。プロモーターの無いコンストラクトで同様の解析をすることで、転写開始（プロモーター）活性ではなく、翻訳開始（IRES）活性であることを確かめている。また、変異導入によって Ribosomal frameshifting の可能性も排除している。彼らはウミシイタケルシフェラーゼの翻訳開始のために EMCV というウイルスの IRES を用いているが、これを除いた場合にも、ホタルルシフェラーゼの翻訳が正常に起こった。すなわち、ホタルルシフェラーゼ（下流 ORF）の翻訳はウミシイタケルシフェラーゼ（上流 ORF）の翻訳とは独立である。これは、SART1 の結果とは全く異なる特徴である。ここまではルシフェラーゼを用いた *in vitro* の解析だが、最後にこれらの結果を踏まえて転移実験を行なっている。興味深いことに、ORF2 の開始コドンの直後に UAA を導入して翻訳を終結させた場合にも、L1 の転移が確認されている。二番目の AUG は転移に必須なエンドヌクレアーゼの内部なので、マウスの L1 の ORF2 の翻訳が、少なくとも一部は、AUG から始まっていないことがわかる。同様に 2 つの IRES を持つ Cricket Paralysis Virus では ORF2 の翻訳開始は AUG からではなく、CCU からであることが示されている（Wilson et al. 2000）。

Li らの解析結果には一つ問題がある。彼らの結果では、ORF1 の IRES よりも ORF2 の IRES の方が活性が強いのである。一般に構造蛋白質である ORF1 蛋白質は機能蛋白質である ORF2 蛋白質よりもはるかに大量に合成される。SART1 では定量はされていないが、ORF1 蛋白質の翻訳量は ORF2 蛋白質の数百倍から数千倍と考えられている。マウスの L1 が彼らの結果どおりの量比で蛋白質を合成すると ORF2 蛋白質の量が膨大になってしまう。彼らは Discussion の中で、ORF1 蛋白質が ORF2 蛋白質の翻訳を阻害する可能性を指摘している。ORF1 蛋白質は RNA と結合する活性があるので、この可能性は確かにある。一方、彼らは 5' キャップ依存的な通常のスキャニングによる ORF1 の翻訳に関して全く解析を行っていないため、ORF1 上流の IRES が翻訳量のわずかな部分にしか参与していない可能性も残っている。

ところで、Alisch らのマウスの L1 の結果と、Li らの結果とでは矛盾がある。Alisch らの結果では、IRES による ORF2 翻訳を否定しているし、Li らの結果では、ORF1 の翻訳は ORF2 の翻訳に影響を与えないことが示されている。それではどちらが正しいのだろうか。細かく見ると、Alisch らは、マウスの L1 については ORF2 翻訳に対する ORF1 の翻訳の影響は調べていない。一方、Li らの解析では、ORF2 は完全にホタルルシフェラーゼに置き換えられてしまっているので、ORF2 内部の配列が ORF2 の翻訳開始に与える影響が失われている。転移実験において、AUG の直後に終止コドン挿入した場合にもある程度の頻度で L1 が転移するのは、Alisch らの結果が示す翻訳機構によるのかもしれない。Alisch

らの解析は主にヒトの L1 についてのもので、Li らの解析は IRES を前提とするものなので、現時点ではどちらが正しいのか判断が付きかねる。今後、ここに焦点を絞った解析がなされることを期待したい。

SART1 の ORF2 翻訳機構が全ての non-LTR レトロトランスポゾンの翻訳機構を説明できないのと同様に、ヒトの L1 の Translational coupling に類似した機構、マウスの L1 の IRES による ORF2 の翻訳も普遍的な機構ではない。これらの研究によってわかった ORF2 翻訳機構は膨大な多様性を持った翻訳機構の氷山の一角なのかもしれない。あるいは ORF2 翻訳機構の種類は限られていても、進化の過程で容易に獲得されうるものなのかもしれない。最後に一つ指摘しておかなければならないことを挙げると、non-LTR レトロトランスポゾンにおいては、ORF1 と ORF2 が in frame でつながったものは見つかっていない。LTR レトロトランスポゾンとは大きく異なる点である。ORF1 と ORF2 とが別々の蛋白質として合成される必然性が何処かにあり、翻訳機構を制限している可能性もある。興味深いことに、Alisch らの実験結果では、L1 の ORF1 と ORF2 を融合させると、転移が確認されなくなる。Non-LTR レトロトランスポゾンにおいては、プロテアーゼによる蛋白質の切断が報告されていないのも意味深である。もちろん、これは LTR レトロトランスポゾンやレトロウイルスの祖先形に近い non-LTR レトロトランスポゾンがプロテアーゼを獲得する前の形態である、ということも一つの説明ではあるのだろう。ただし、細胞の持つプロテアーゼを使用することも可能なのだからそれで全てが説明できるわけではない。Non-LTR レトロトランスポゾンがより単純な形態をしているからと言って、全てを LTR レトロトランスポゾンやレトロウイルスのアナロジーで考えてはいけないということを上述のような研究は教えてくれている。

Kojima KK, Matsumoto T, Fujiwara H.

Eukaryotic translational coupling in UAAUG stop-start codons for the bicistronic RNA translation of the non-long terminal repeat retrotransposon SART1.

Mol Cell Biol. 2005 Sep;25(17):7675-86.

Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV.

Unconventional translation of mammalian LINE-1 retrotransposons.

Genes Dev. 2006 Jan 15;20(2):210-224.

Li PW, Li J, Timmerman SL, Krushel LA, Martin SL.

The dicistronic RNA from the mouse LINE-1 retrotransposon contains an internal ribosome entry site upstream of each ORF: implications for retrotransposition.

Nucleic Acids Res. 2006 Feb 6;34(3):853-864.

Wilson JE, Powell MJ, Hoover SE, Sarnow P.

Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites.

Mol Cell Biol. 2000 Jul;20(14):4990-4999.

2006/03/07

2006/03/10 加筆修正

小島 健司 著

禁 無断複写転載