

出“X”記あるいは出“精巢”記：retrogene

「L1 Gives You Wings : Alu と processed pseudogene」では、mRNA が L1 の蛋白質によって逆転写され、イントロンを持たない processed pseudogene ができることを紹介した。L1 自身とは異なり、mRNA の配列は内部にプロモータを持っていないので転移先では通常転写されない。しかし、転移先に元々あった遺伝子のプロモータを流用するなど、何らかの機構で転写されれば、遺伝子として機能することができるはずである。ここでは、mRNA の逆転写を介して転移した配列を retrocopy、その内機能を持っていると考えられるものを retrogene と呼ぶことにする。従って retrogene は processed gene と同義で、retrocopy から retrogene を除いたものが processed pseudogene である。

マウスの *Utp14b* は必須機能を持った retrogene として哺乳類で最初に見つかった遺伝子である (Bradley et al. 2004; Rohozinski and Bishop 2004)。マウスの *jsd* (juvenile spermatogonial depletion、若年性精子幹細胞枯渇) 変異体は精子形成に異常が生じる変異体である。Bradley らは *jsd* 変異体の原因遺伝子の位置をポジショナルクローニングにより第 1 染色体上の 272 kb の領域まで絞り込んだ

(Bradley et al. 2004)。この 272 kb にコードされる遺伝子を予測し、変異体と野生型とで配列を比較したところ、*Utp14b* の GG が CTTTTC に変化していることが *jsd* 変異の原因であることがわかった。この変異はフレームシフトにより *Utp14b* 蛋白質の C 末側の欠失を引き起こす。*Utp14b* は別の遺伝子 *Acsl3*

(Acyl-CoA synthetase long-chain family member 3) の第 2 イントロン内に蛋白質コード領域があり、上流の 2 つのエクソンを *Acsl3* と共有していた。*Utp14b* と *Acsl3* は共に精巢で転写が見られる。*jsd* 変異体でも *Acsl3* と *Utp14b* の転写、スプライシングに異常は見られないことから、*Utp14b* の蛋白質の C 末側の欠失が *jsd* 変異の原因であると Bradley らは結論づけた。

一方、Rohozinski らはポジショナルクローニングによって絞り込んだ 1.5Mb の領域にある精巢での転写産物を解析し、1 つの転写産物が、*jsd* 変異体では見られないことを発見した。これが *Utp14b* であった。*Utp14b* は別の遺伝子 *Facl3* (fatty acid CoA ligase long-chain 3) の第 1 イントロン内にある 3 つのエクソンで構成され、3 番目のエクソンに蛋白質コード領域がある。Rohozinski らは、*Utp14b* の全長を含み、*Facl3* や他の遺伝子の全長は含まない BAC を *jsd* 変異体マウスに導入し、精子が正常に形成されることを確認している。以上から、*Utp14b* の蛋白質の C 末側の欠失が *jsd* 変異の原因であると結論づけた。後に明らかになる (Zhao et al. 2007) ことだが、*Acsl3* と *Facl3* は同じ遺伝子であり、選択的スプライシングにより 2 種類の mRNA が転写される。Bradley らの言うところの第 2 エクソンは選択的スプライシングによって除かれる場合があるため、*Utp14b* は

Rohozinski らの論文では第 1 イントロン内にあるとされた。以後は Bradley らの *Acs13* の表記に従うことにする。更に、Bradley らと Rohozinski らの発見した *Utp14b* の転写産物も 5'側のエクソンが異なっている。

Utp14b の蛋白質コード領域は X 染色体上にコードされる *Utp14a* と 77%の塩基配列相同性を持つ。また、*Utp14a* は蛋白質コード領域に 14 個のイントロンを持つのに対して、*Utp14b* は蛋白質コード領域にイントロンを持たない。このことから、*Utp14b* は *Utp14a* の mRNA が逆転写されて *Acs13* のイントロンに挿入され、(Bradley らの研究によれば少なくとも一部の mRNA については) *Acs13* の転写機構を借りて転写、機能するようになったと考えることができる。オリジナルの遺伝子 *Utp14a* は酵母の *Utp14* のオルソログで、rRNA のプロセシングに必須の蛋白質をコードし、幅広い組織で発現する。一方の *Utp14b* は精巣と脳でのみ発現し、精子形成に必須の機能を持つ。ヒトには *Utp14b* は存在しないが、*Utp14a* から別に転移した *UTP14C* という遺伝子が存在する。Bradley らは 16 種の有胎盤哺乳類を調べ、13 種から *Utp14a* に由来する retrocopy を発見している。これらの遺伝子の蛋白質コード領域には変異が蓄積されておらず、しかも、アミノ酸を変えない塩基置換が多く見られる。配列比較からは、これらの配列は独立に *Utp14a* の mRNA が転移して生まれたと考えられる。これらの retrogene は、起源は異なりながらも同じ機能を果たしているのかもしれない。

その後の解析で、*Utp14b* には 5 種類の転写産物があることが明らかとなった (Zhao et al. 2007)。1 つは *Acs13* とエクソン 2 つを共有し、精巣と脳で強く発現するが、他の組織でも弱く発現している。これが Bradley らの見つけた転写産物である。*Acs13* の発現も精巣と脳で強く、他の組織で弱く発現することからも、*Acs13* と *Utp14b* のこの転写産物も同じプロモータから同じ蛋白質群によって転写が開始されていることがわかる。ただし、*Acs13* が精巣よりも脳でより強く発現するのに対して、*Utp14b* は精巣でより強く発現が観察される。これは mRNA の安定性やスプライシングの効率など、転写後の制御機構が働いていることを示唆する結果である。一方、残り 4 つの転写産物は *Acs13* の第 1 イントロン内部に 2 つのエクソンを持ち、その選択的スプライシングによって合成される。この内の 1 つが Rohozinski らの記述した転写産物である。前側のエクソン 1b には内部にスプライシングドナー配列があるため、2 番目のエクソン 1.5 の有無と 1b の長短によって 4 通りの配列が生まれる。この 4 つはどれも *Acs13* の第 2 エクソンを含むので、*Acs13* の第 2 エクソン以降の配列は 5 つの転写産物に共通である。この 4 種類の転写産物は精巣特異的に、より具体的には精巣の生殖細胞特異的に転写される。この 4 つの転写を制御するプロモータは *Acs13* の第 1 イントロン内部にあると考えられるがどの配列が機能しているのかはわかっていない。

Utp14b のどの転写産物もパキテン期の精母細胞で発現している。一方、*Utp14a* はパキテン期の精母細胞ではほとんど、あるいは全く転写されていないことも明らかとなった。パキテン期には X 染色体が凝集し、ほとんど転写されない状態になることが知られている。この X 染色体の凝集によって *Utp14a* の転写は確かに抑制され、代わりに *Utp14b* が発現して相補していると考えられることができる。

興味深いことに retrogene には *Utp14b* のように、X 染色体上の遺伝子に由来し、精巢で発現するものが多い (Emerson et al. 2004)。Emerson らは、ヒトゲノムで同定されている遺伝子同士を比較することで、655 個のオリジナル遺伝子—retrogene 対を発見した。内、366 個は異なる染色体上に位置していた。更ここから、retrogene と元の遺伝子とを比較した際の、非同義置換率の同義置換率に対する割合 Ka/Ks が 0.5 未満で、かつ転写されていることを示すデータがあるものを選び出すと 94 対が残った。同様の解析により、マウスゲノムからも 105 対が見つかった。

retrogene の元となった遺伝子の乗っている染色体を調べると、X 染色体だけが突出して多かった。この傾向はヒトでもマウスでも違いは無く、染色体の長さを考慮した場合でも非常に高い割合であった。これに対して、変異の蓄積している retrocopy、すなわち processed pseudogene について同様の解析を行なうと、X 染色体由来の processed pseudogene が有意に多いわけではなかった。また、精巢や卵巣での遺伝子発現についても X 染色体上の遺伝子の発現が多いわけではない。これらの事実は、retrocopy が X 染色体から常染色体に転移する確率は他の染色体の場合と変わらないのに、retrocopy が偽遺伝子化する確率が低い（遺伝子化する確率が高い）ことを示している。この原因として、オスの減数分裂期に起こる X 染色体の凝集により X 染色体上の遺伝子の発現が抑制される現象が考えられた。そこで EST を利用して発現を調べると、X 染色体上の遺伝子に由来する retrogene では 10/13 が精巢で発現するのに対して、他の retrogene では 20/45 しか精巢で発現していなかった。以上の結果からは、X 染色体上の遺伝子は精巢で発現が阻害されるため、常染色体に転移した retrogene が精巢で発現することでその機能を代替する、という現象が進化上有利であることが示唆される。

X 染色体は元の遺伝子が乗っている確率が高いのに加えて、retrogene が乗っている確率も高い。常染色体と比べて有意に多くの retrogene が転移して来ている。この一因は X 染色体では組み換えが起こりにくく挿入が蓄積しやすいことであるが、それだけでは違いを説明しきれないことが統計検定からは示された。残りの違いは、オスでは数が多いことで有利になり、メスで不利になるような遺伝子が、メスで発現が抑制される X 染色体上に転移した結果かもしれないと Emerson らは考察している。実際、X 染色体上の retrogene では、1/7 しかメスの

組織で発現が観察されないのに対して、常染色体上の retrogene では、32/45 がメスで発現する。

Marques らは retrocopy の遺伝子化の頻度を解析した (Marques et al. 2005) 。ヒトのゲノム配列から 3951 個の retrocopy を見つけ、その中の 705 個がオリジナル遺伝子の蛋白質コード領域に相当する領域に終止コドンやフレームシフト変異を含んでいないことを示している。これら”intact”な retrocopy の転移時期は、霊長類の進化の全期にまたがっているが、3,800-5,000 万年前にピークを持っている。同義置換と非同義置換との比較からは、705 個の内、57-76 程度が遺伝子として働いていることが示唆された。従って 6,300 万年ほどの霊長類進化の過程で 100 万年に 1 個程度の頻度で retrogene が出来ている計算になる。intact な retrocopy は変異を含んでいる retrocopy と比較すると、同義置換に対する非同義置換の比率が低い場合 ($Ka/Ks < 0.4$) に多い。これは、負の淘汰圧を受けていることを示唆している。

続いて Marques らは intact な retrocopy から同義置換率が低い、すなわち比較的新しい 38 個を、 Ka/Ks に関係なく選び、8 種の哺乳類で有無を調べた。5 つは旧世界猿に共通し、5 つは類人猿特異的、6 つはヒト特異的、などと転移時期の異なる retrogene が選ばれている。更に、600 塩基以上でヒト・チンパンジー・ゴリラ全てに存在する 23 個の retrocopy の配列を比較すると 8 つだけが 3 種全てで intact であった。この内 1 つは既に精巣での蛋白質発現の報告のある *PABP3* である。シミュレーションの結果からは、残り 7 つの遺伝子は進化的に中立に維持されているとは考えにくく、遺伝子として淘汰を受けている retrogene であることが示唆された。この 7 つの retrogene はどれも精巣で強く転写され、内 3 つではオリジナルの遺伝子が X 染色体上に乗っている。

Vinckenbosch らは retrocopy の遺伝子化と転写との関係を調べている (Vinckenbosch et al. 2006) 。3,590 個の retrocopy を見つけ、EST と全長 cDNA の情報を利用して retrocopy の転写を解析した。約 1/3、1,080 個の retrocopy について少なくとも 1 つの EST が見つかった。また、intact な retrocopy の転写されている割合は、retrocopy 全体における転写されている割合に比べて 1.7 倍と多かった。また、EST の数が多い retrocopy では、intact である可能性が高く、 Ka/Ks の平均も低かった。以上の結果は、よく転写されている retrocopy には機能を持った遺伝子が多いことを示唆している。

retrocopy の挿入位置と転写との関係を調べると、イントロンに挿入されている retrocopy では 42% が転写されているのに対して、遺伝子間領域に挿入されている retrocopy では 25% しか転写されていない。また、遺伝子間領域の retrocopy では転写されているものの方が転写されていないものに比べて遺伝子に近い傾向があった。イントロン内に挿入された retrocopy の転写が挿入位置の遺伝子の

転写に依存しているとするならば、その遺伝子は retrocopy と融合した mRNA を作っている可能性がある。実際、36 個の融合 mRNA を見つけることができた。内 19 個では、蛋白質をコードしていないエクソンと retrogene が融合しており、残り 17 個では、蛋白質をコードするエクソンと retrogene とが融合していた。また、EST の数から見た転写量は、融合 mRNA の場合には、retrogene 単独の場合よりも有意に多かった。元々あった遺伝子との融合以外にも 27 個の retrocopy が新しくエクソンを獲得していた。新しいエクソンを獲得した retrocopy は転写量の多いものに特に多い。

転写される器官は精巣が多いことは先述した。精巣の EST の量は、retrocopy は通常の遺伝子の 2 倍以上で、intact な retrocopy は intact でない retrocopy の 2 倍以上であった。興味深いことに新しい retrocopy の EST は数が少ないが精巣に偏っており、マウスにも存在するような古い retrocopy の EST は数が多く幅広い組織に分布していた。これは、精巣で転写される内に retrocopy は機能を持つ retrogene へと進化し、後に幅広い組織で発現できるように制御領域を獲得するという一般的な方向性があることを示唆している。

もちろん、新しい遺伝子ができるばかりが retrocopy の挿入の結果ではない。挿入された側の遺伝子の転写は、retrocopy の配列によって阻害される可能性がある。実際、高頻度に転写されている retrocopy においては、イントロンに挿入されているものの割合は低い。また、挿入された側の遺伝子の転写と同じ向きに挿入される retrocopy の割合が低いことは、sense 鎖の配列が転写されるとプロセシングに悪影響が出ることを示唆している。

TACSTD2 は失明する病気 gelatinous drop-like corneal dystrophy（膠様滴状角膜変性症）の原因遺伝子であることが報告されている。この *TACSTD2* はイントロンを 8 つ持つ *TACSTD1* の retrogene であることも Vinckenbosch らの解析から明らかとなった。

グルタミン酸脱水素酵素をコードする *GLUD2* も retrogene である (Burki and Kaessmann 2004)。*GLUD2* はイントロンを持つ *GLUD1* の mRNA が転移してきた新規遺伝子である。Burki らは各種霊長類で *GLUD2* の有無を調べ、類人猿（ヒト、チンパンジー、ゴリラ、オランウータン、テナガザル）では *GLUD2* と target site duplication が見つかるが、旧世界猿では挿入前の状態であることを明らかにした。従って *GLUD2* は旧世界猿と分岐した後、類人猿で転移した retrogene である。*Utp14b* は精巣と脳で発現し、精巣で必須の機能を担っていた。逆に *GLUD2* は精巣と脳、網膜で発現し、脳で重要な役割を持っている。*GLUD2* は *GLUD1* に比べ、高濃度の GTP 存在下で高い活性を持ち、アロステリックな活性化因子である ADP やロイシンの濃度に依存して 25 倍以上に活性化される。また、比較的低い pH でも活性を持つ。これらの特徴は神経伝達物質の一種である

グルタミン酸を素早く用意するのに適している。このような有利な特徴は GLUD2 特異的なアミノ酸置換によって達成されており、例えば、456 番目のアラニンは高濃度の GTP 存在下での活性に重要であり、443 番目のセリンは活性を低下させ、ADP 存在下で大きく活性化の程度を上昇させるのに寄与している。これらのアミノ酸は類人猿の GLUD2 に共通しており、類人猿の共通祖先で正の淘汰圧による加速進化が起こっていたことが解析から支持された。言い換えると、GLUD2 は転移後短期間の内に脳で効率よく機能するように進化し、類人猿の脳の機能向上に働いている可能性が示されたと言える。

このように見てくると、精巢で発現が抑制される X 染色体由来の遺伝子の機能が、常染色体に飛んだ retrocopy が遺伝子 (retrogene) 化することで相補され、後にこの retrogene が新しい機能を獲得するという流れが、哺乳類での新規遺伝子の誕生に寄与しているとわかる。遺伝子化の最初のステップは別の遺伝子の転写機構に乗かって転写される、つまり選択的スプライシングの 1 多型になることであるが、やがて自身のプロモーターを獲得して自立した遺伝子となっていく。個人的にはこのような進化的機構がいつ誕生し発展してきたのかが興味深い。なぜなら哺乳類で retrocopy を転移させるのはほとんどが L1 であり、この L1 の性質は non-LTR retrotransposon 全般から見るとかなり特殊だからだ。XY タイプの性決定と性染色体も哺乳類の進化の過程で生まれてきた物である。もしかするとこのような retrogene の生成機構も哺乳類に独特のものかもしれない。

Bradley J, Baltus A, Skaletsky H, Royce-Tolland M, Dewar K, Page DC.

An X-to-autosome retrogene is required for spermatogenesis in mice.

Nat Genet. 2004 Aug;36(8):872-6.

Rohozinski J, Bishop CE.

The mouse juvenile spermatogonial depletion (jsd) phenotype is due to a mutation in the X-derived retrogene, mUtp14b.

Proc Natl Acad Sci U S A. 2004 Aug 10;101(32):11695-700.

Zhao M, Rohozinski J, Sharma M, Ju J, Braun RE, Bishop CE, Meistrich ML.

Utp14b: a unique retrogene within a gene that has acquired multiple promoters and a specific function in spermatogenesis.

Dev Biol. 2007 Apr 15;304(2):848-59.

Emerson JJ, Kaessmann H, Betrán E, Long M.

Extensive gene traffic on the mammalian X chromosome.

Science. 2004 Jan 23;303(5657):537-40.

Marques AC, Dupanloup I, Vinckenbosch N, Reymond A, Kaessmann H.

Emergence of young human genes after a burst of retroposition in primates.

PLoS Biol. 2005 Nov;3(11):e357.

Vinckenbosch N, Dupanloup I, Kaessmann H.

Evolutionary fate of retroposed gene copies in the human genome.

Proc Natl Acad Sci U S A. 2006 Feb 28;103(9):3220-5.

Burki F, Kaessmann H.

Birth and adaptive evolution of a hominoid gene that supports high neurotransmitter flux.

Nat Genet. 2004 Oct;36(10):1061-3.

2011/02/10

小島 健司 著

禁 無断複写転載