



琉球大学

University of the Ryukyus

Title	データベースを利用したヒト内在性レトロウイルス関連遺伝子が持つ転写能の検討
Author(s)	小田, 高也
Citation	
Issue Date	2006-03
URL	http://ir.lib.u-ryukyu.ac.jp/handle/123456789/314
Rights	

琉球大学学術リポジトリ
University of the Ryukyus Repository



琉球大学学術リポジトリ
University of the Ryukyus Repository



データベースを利用したヒト内在性レトロウイルス
関連遺伝子が持つ転写能の検討

16570004

平成16年度～平成17年度 科学研究費補助金
(基盤研究 (C)) 研究成果報告書

平成18年3月

研究代表者 小田 高也
琉球大学 大学院医学研究科 助教授

はしがき

平成15年4月に公表された10年余にわたる国際プロジェクトの成果により Transposable Element (TE) はヒトゲノムの約45%という膨大な領域を占めていることが明らかとなった。同時にかつては10万とも予想されていたヒトの遺伝子は、はるかに少ない“数万個”しか存在しないらしいことも分かってきた。ヒトはショウジョウバエが持つ僅か2倍あまりの遺伝子で10万種を越える酵素やホルモンなどのタンパク質を作り分けており、更には言語を操るといった能力に代表される脳の高次機能を実現している事実は注目に値する。「少ない遺伝子が複雑な生命活動を支える仕組みを解明すること」は生命科学の今後の重要な課題の1つともいえよう。そんな中、かつて存在意義が不明なため意味をなさない存在“ジャンク”と考えられていた TE は、遺伝子の発現調節との関与やクロマチン構築およびヒトの進化における役割という観点から関心が高まってきている。

TE の一つであるヒト内在性レトロウイルス (HERV: Human Endogenous Retrovirus) はヒトゲノムの8%の領域を占めており、L1 に代表される LINE (long interspersed nucleotide element) や Alu に代表される SINE (short interspersed nucleotide element) に比べてコピー数が少ないグループを形成している。LTR-gag-pol-env-LTR という構造が暗示する如く、転写調節・自己複製・組込という過程を通して数を増やし現在に至ったと考えられている。ヒトゲノムに進入した時期は多岐にわたり、新しいものではヒトとチンパンジーが分岐した後に組み込まれたと考えられるコピーの報告もある。“組込”は時として宿主にとって致命的となりうるプロセスであるが霊長類は巧みにそれを制御してきた。事実ゲノム中に散在する大多数の HERV は転写活性能を失い、自己複製に必要とされる蛋白コード領域には致命的な変異が生じていてあたかも“死んで”いるかの様に見える。しかし、中には依然として活発に活動し臓器特異性を示す転写能を保持しているものも報告されている。一例を挙げると、唾液腺で発現する Amylase 遺伝子の転写制御に HERV の LTR が関与していることが明らかにされたように、これら転写物は宿主に対して有益なものとして存在し宿主が意図して“生き残らせている”様に思われる。これらの機能遺伝子を系統的に単離しその機能を明らかにしてゆくことは、HERV が持つヒトの進化に果たした役割とその機能異常としての疾病の病態理解に役立つ可能性があると考えられる。

研究組織

研究代表者：小田 高也（琉球大学大学院医学研究科助教授）

交付決定額（配分額）

（金額単位：円）

	直接経費	間接経費	合計
平成16年度	1,700,000	0	1,700,000
平成17年度	1,300,000	0	1,300,000
総計	3,000,000	0	3,000,000

研究発表

（1）学会誌等

- Okahara G, Matsubara S, Oda T, Sugimoto J, Jinno Y, Kanaya F.
Expression analyses of human endogenous retroviruses (HERVs): tissue-specific and developmental stage-dependent expression of HERVs.
Genomics, 84(6):982-990, 2004
- 小田高也, ヒト内在性レトロウイルスと統合失調症. 脳と精神の
医学, 16・2:87-93, 2005

（2）口頭発表

発表者名：小田 高也

演題名：脳での発現能を有するヒト内在性レトロウイルス関連遺伝子の探索

学会名：第13回日本精神・行動遺伝医学会

場所：九州大学同窓会館（福岡）

年月日：平成17年10月1日

（3）出版物 なし

研究成果による工業所有権の出願・取得状況：なし

研究成果

1. 研究の背景

ゲノムプロジェクトの進行に伴って、かつて意味をなさない存在“ジャンク”と考えられていた TE が注目を集める中で、以前から TE の 1 グループである HERV とヒト疾患との関与を示唆する報告がなされていた。例を挙げると I 型糖尿病患者膵臓培養上清中より単離された HERVK がスーパー抗原活性を持つことが示されたこと (Conrad B.1997), 多発性硬化症患者髄液中から HERVW が検出されたこと (Komurian-Pradel F.1999) など自己免疫疾患と HERV の関連を示唆する報告がある。また, Syncytin と名付けられた別の HERVW が胎盤での合胞体形成に関与するという HERV が提供するヒトの生理機能の関与を示唆する報告もある (Mi S.2000)。近年, 統合失調症および分裂感情障害症例の脳脊髄液および血漿中で HERV 関連の転写物が検出されたことから, 通常は脳組織での発現が抑制されている HERV とこれらの機能性精神疾患発症との関連も注目されている (Karlsson H.2001)。当教室では互いに配列の相同性が高い HERV の発現をローカス特異的に検証する独自の方法を考案し, HERVK, HERVE, HERVH のサブファミリーより活性型の HERV を同定し報告してきている (Sugimoto J.2001, Shiroma T.2001, Nakamura A.2003)。

2. 目的

dbEST や nr 等の公共データベースを利用する事により, HERV の中でも転写活性を保っているものを効率的に探索する。そして, 近年飛躍的に充実してきている Bioinformatics を利用して, 同定された HERV ローカスに隣接する遺伝子を見出しその遺伝子の発現制御に HERV の発現が関与しているかどうかの検討もあわせて行う。更には, 霊長類 DNA サンプルを用いてその HERV が進化の過程のどの時期からヒトゲノムと共生するようになったかを見極めることにより, 進化に果たした役割を考察する事を目的とする。

3. 方法

① Web-BLAST による HERV 関連 data entry の収集

2000 年に Tristem M. が HERV を 22 種類に分類した際に用いた代表的な配列と、Rebase (www.girinst.org/) に HERV 関連配列として登録されているものを加えた延べ 96 の塩基配列を Query として dbEST を Web-BLAST (www.ncbi.nlm.nih.gov/BLAST/) で検索した。検索オプションの e-value は検索漏れを防ぐために default 値の 10 よりも大きくして 50 (ないし 100) を用いた。完全長 cDNA プロジェクト等の進展により mRNA 配列として nr データベースに登録されている配列数が飛躍的に増加していることも考慮して、同じ Query で nr データベースも検索した。ただし、検索オプションの “Limit by entrez query” に “biomol_mma[PROP]” を用いることにより HTGS entry 等の混入を防ぐように配慮した。予備実験で Web-blast で得られた EST entry の由来臓器が多数に及ぶ場合、それ以降の解析 (特に RT-PCR の template の選択など) で判断が困難な場合が予想された。よって、まずは脳組織を対象を絞って解析を行った。具体的には、“Limit by entrez query” に qualifier で定義した brain 関連の term をブール論理で組み合わせて入力した。

② in-house BLAST による mapping

従来、HERV 構造の中でも 3'LTR は EST の 3' entry に含まれることが多く、同一サブファミリー内では LTR 配列の相同性が極めて高いことと、各 EST の配列の正確さに 100% を期待できないことから mapping を含めた解析が困難であった。しかし、ヒトゲノム配列が利用できるようになった現在、BLAST によりヒトゲノム中からその entry が存在する確率が高いローカスを推定することが可能になっている。

まず、96 個の Query より得られた BLAST の結果は “-b 0” オプションにより Alignments 表示を省きデータの簡素化を図り、Descriptions 部分のテキストを GREP や PERL 等のシェルスクリプトで処理して accession number を抽出する。Batch Entrez (www.ncbi.nlm.nih.gov/entrez/batchentrez.cgi?db=Nucleotide) を利用して対応する核酸塩基配列を FASTA format で一括ダウンロードする。

BLAST プログラムは各プラットフォーム毎に local で実行可能なファイル形式のものが入手可能で、local で稼働させることによりネットのトラフィックやホストの

負荷に依存する Web-base の BLAST よりも安定した持続的検索が可能になる。そこで、NCBI の ftp サイトより PowerPC 用のバイナリファイル(blast-2.2.10)をダウンロードして PowerMacG5, 2.0GHz (dual)にセットアップした。また、build_35.1 のヒトゲノムデータも ftp サイトよりダウンロードして BLAST で使用できるように formatdb した。次に、FASTA format の配列を query として in-house blast を行うことによりその entry の mapping を行った。

③ データ解析

BLAST 結果を表示する Format の Alignment view を default の Pairwise でなくて Hit Table を選ぶことにより、“query id, subject ids, % identity, alignment length, mismatches, gap opens, q. start, q. end, s. start, s. end, evaluate, bit score”の各項目が得られる。24 本の染色体に相当する “subject ids”と “s. start”の 2 項目を指標にデータにソートをかける (Microsoft Excel を使用)。更に、Query に用いた HERV のサイズが最大約 9kb であることを考慮して、近傍 10kb 以内に他の entry が mapping されたデータにマークをつけるよう表示関数を指定した。これにより、どのローカスに HERV 関連の entry が集簇しているかが容易に判別可能となった。

④ RT-PCR & Sequence

Mapping により明らかとなったローカスに関して、まず UCSC の Human Genome Browser Gateway (genome.ucsc.edu/cgi-bin/hgGateway)にて収集したゲノム情報を参考にしてそのローカスに HERV の存在を確認した後に、その転写物の存在をベンチワークで確認するために RT-PCR を行った。Template には、胎児脳由来の total RNA を使用した。リピート配列を研究対象とする場合に常に注意しなければならないポイントとして、極めて高い相同性を示す配列が複数コピー (サブファミリーによっては数千) ゲノム中に存在することを考慮して、プライマーは mRNA 中の同じサブファミリーのメンバーを偏りなく増幅させるために (RNA 中のサブファミリーメンバーの mRNA の population がクローン化産物の population に反映されるために) degenerate プライマーを使用した。プライマーによる増幅に bias がかかっているかどうかの評価は、genomic DNA を PCR で増幅した際に得られたクローン間の配列に片寄りがないことで確認した。RT-PCR, genomic PCR 共に 3~4 サイクル毎にサンプリングを行い、増幅が plateau に達する前の PCR product を使用した。PCR product は電気泳

動で対応するバンドを確認した後 agarose gel より切り出し, GeneClean で精製した後 pGEM-T easy vector でクローニングした. Sequence には Applied Biosystems の ABI PRISM[®] 310 Genetic Analyser を用いた. クローン数にして 30 個前後の配列比較を行い, 予測したローカスでの発現の程度を判断した.

⑤ 霊長類ゲノムへの進入時期の検討

胎児脳由来 RNA を用いた RT-PCR & Sequence で検証できた 2 つのローカス (1p36.13 の HUERS-P3b に属する HERV 転写物および 5p15.33 の HERVK9 に属する転写物) に加え, ノーザン解析にて胎盤での特異的発現が確認された 21 番染色体の 1 ローカスを加えた合計 3 ローカスの HERV に関して, 霊長類の進化との関連を考察するためにゲノム DNA を用いて宿主への進入時期の検証を PCR ベースの実験で行った. de Parseval らの報告を参考に, Genome Browser から得られたゲノム配列を元にして 5' および 3' の各 LTR 内とその flanking 領域に位置するプライマーセットを作成した. ヒト genomic DNA で PCR 条件を検討した後, 平成 17 年度に採択された京都大学霊長類研究所共同利用研究員として愛知県犬山市の同施設に赴き, 所内対応者である景山節教授よりチンパンジー (雌雄各 1 サンプル) ・ニホンザル ・アカゲザル ・コモンマーモセット ・フサオマキザル ・ヨザル ・クモザル ・ギャラコを供与していただき, ヒトを含めた 10 サンプルで PCR を行った.

4. 結果

① Web-BLAST による HERV 関連 entry の収集

96 個の塩基配列 (744 - 8,203 nt) を Query として Web-blast を行ったところ, 1,332 個 (延べ 3,142 個) の脳組織関連 EST entry が得られた. accession number を抽出して Batch Entrez で FASTA 形式でファイルを入手したところ 67 - 2,049 nt のサイズであった.

② in-house BLAST による mapping

Web-blast で得られた 1,332 個の entry のローカスを予測するために行った in-house blast では, 結果の “alignment length” の項目に示されたように Query 配列のうち 67 - 2,044 nt にもっとの高い SCORE を示す相同性を示す部位が見いだされ, その “%

identity”は 83-100 %であった. 表 1 に示すように全体の 8 割以上の entry が比較された範囲に関して build_35.1 のヒトゲノム配列との一致率 96%以上を示した.

表 1 : EST entry とヒトゲノム配列(build_35.1)との一致率の内訳

% identity (Id)	entry no.
Id<90	28 (2.1%)
90≤Id<92	51 (3.8%)
92≤Id<94	56 (4.2%)
94≤Id<96	109 (8.2%)
96≤Id<98	284 (21.3%)
98≤Id<100	596 (44.7%)
Id=100	208 (15.6%)
計	1,332

(括弧内は全体に対するパーセントを表す)

③ データ解析

表 2 に entry の染色体毎の分布を示す. 染色体 1 本あたりの平均値は 56 個であるが, 染色体 1 番, 8 番, 19 番などは平均の 2 倍以上の数の entry が属するように思われた. そこで, mapping に使用した build_35.1 の各染色体の総塩基数に対する entry 数の割合を見たところ, 染色体 1 番の値を 1 として比較すると染色体 19 番の 4.47 を除いてすべて 1.5 以内となり, 19 番染色体は HERV 関連転写物に関しては特別な存在であることが示唆された.

表 2 : EST entry の染色体の分布

染色体	Entry no.	染色体	Entry no.	染色体	Entry no.
1	135(8)	9	29(1)	17	19(1)
2	63(8)	10	50(5)	18	12(0)
3	66(7)	11	36(3)	19	157(14)
4	62(8)	12	87(7)	20	31(1)
5	37(2)	13	79(5)	21	20(3)
6	24(3)	14	88(5)	22	16(1)
7	53(5)	15	16(0)	X	54(5)
8	120(3)	16	66(6)	Y	12(1)

(括弧内は 3 つ以上の EST が集簇するスポット数を示す)

また, entry の集簇の程度を Microsoft Excel を用いて解析した結果, 3 つ以上の EST で構成されるスポットが 102 箇所見つかった. 染色体一本あたりの平均は 4.3 個であり, やはり染色体 19 番は 3 倍近くの数のスポットを持っていた. 各染色体の総塩基数を考慮しても 19 番染色体は 6.73 と最大値 (染色体 1 番の割合を 1 として計算. 平均値は 1.13 で 19 番以外は全て 2.10 以下) を示した.

④ RT-PCR & Sequence

解析により見つかった 102 箇所のスポットのゲノム情報を Genome Browser より入手し, 脳組織以外の EST の関与も再検索して候補を絞ったところ, 1p36.13 にある HUERS-P3b と 5p15.33 の HERVK9 が解析候補になった (図 1).

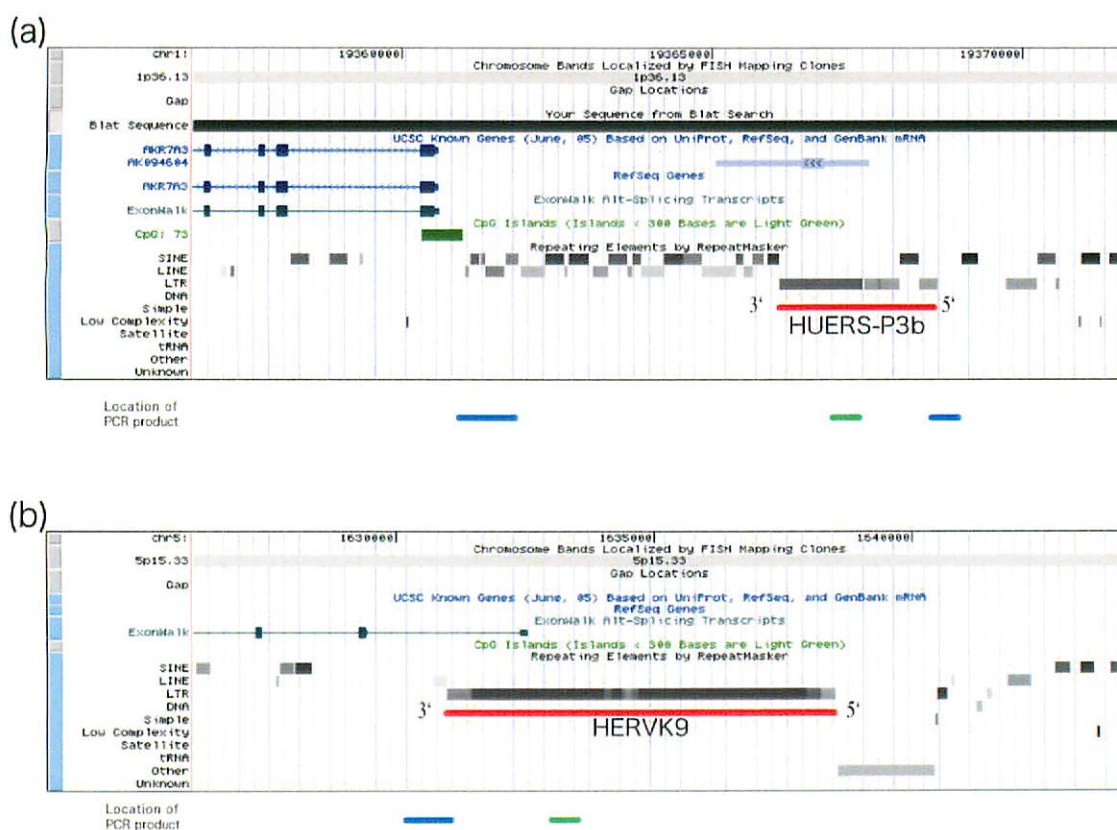


図 1 : (a)1p36.13 の HUERS-P3b と (b)5p15.33 の HERVK9 のゲノム上の概略を示す. 枠外に示すグリーンの線は degenerate プライマーで増幅した領域, ブルーの線は霊長類ゲノム DNA を増幅した領域を示す.

プライマーセット① (GTCCTCCRAARCCARAWTTT , ATGGAGGAAGTGTGCTRRWT) およびセット② (TAATGGTAAADCTTCTTAYT, ATCTGTATGAAATCRTAAYT) を, それぞれ 1p36.13 と 5p15.33 の解析用に合成し RT-PCR を行ったところ, セット①からのクローン化産物は 28 個中 12 個(43%)が 1p36.13 ローカスの配列と同一であり, セット②では 33 クローン中 23 個(70%)が 5p15.33 と一致した. Genomic PCR では各 30 クローンの配列を比較して, プライマーによる増幅効率に bias は認めなかった.

⑤ 霊長類ゲノムへの進入時期の検討

1p36.13 の 5' 用にセット③ (TTTTACAGGCTGCACTTCGTTAG , TAGCCTTTTTCTCTGTGTAGC), 3'用にセット④ (CCGTGTGGTCCTGGAGAATG, TTGAGACAGGGTCTCAATCTGCAG) , そして 5p15.33 用にセット⑤ (TCCTCCTCCTTGCTTCTCCCTCTA, CATGCTGGAAGATTGTCGGTTTA) を作成した. 更に, ノーザン解析で胎盤での特異的発現を認めた 21q22.3 の HERVF(typeB):FB1 (図2) 用にセット⑥ (CAGGTGCTGACCTAAAGGAAGAAG, CTGAAAGCTTGCGGGATGAATAAC)を準備して, 霊長類ゲノム DNA を用いて PCR を行った.

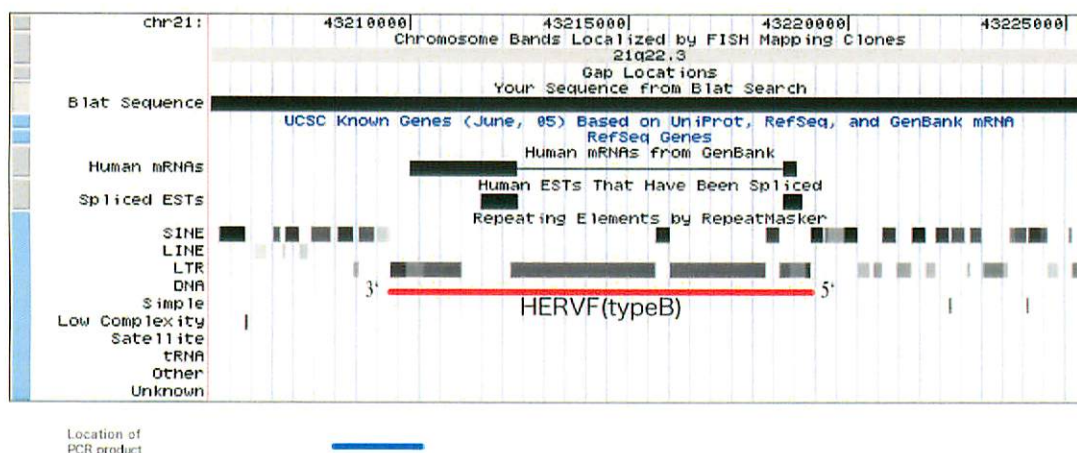


図2 : 21q22.3 の HERVF(typeB)のゲノム上の概略を示す. 枠外に示すブルーの線は霊長類ゲノム DNA を増幅した領域を示す.

結果を表3に示す。1p36.13 の 3'側, 5p15.33 の 5'および 3'の2カ所ともヒトおよびチンパンジーでのみ増幅産物が得られた。21q22.3 の 3'側ではヒトでのみ増幅産物が得られた。

表3：霊長類ゲノムDNA を利用したHERV 進入時期の検討

		チンパンジー(♂)	チンパンジー(♀)	ニホンザル	アカゲザル	ゴモンマーモセット	フサオマキザル	ヨザル	クモザル	ギヤラク	ヒト
HUERS-P3b	3'LTR 領域	○	○								○
HERVK9	5'LTR 領域	○	○								○
	3'LTR 領域	○	○								○
HERVF(typeβ)	3'LTR 領域										○

(○印はヒトと同じサイズのPCR 産物が得られたことを示す)

5. 考察

HERV の存在意義およびヒトの疾患との関連を明らかにする目的で、転写活性を保持する HERV 関連転写物を系統的に単離同定するために、mRNA に由来する核酸の公共データベースを BLAST プログラムでスクリーニングした。ベンチワークでの解析対象を見つけ出すというコンピューター上の作業はヒトゲノムのアノテーション作業が進行していることも助けとなり順調に進んだ。しかし、反復配列を検証するために degenerate プライマーを使用することが必須だったため、その調整も含めたベンチワークに予想以上に手間がかかり、当初予定していたノーザン解析や Open Reading Frame の確認のための in vitro Transcription / Translation assay を含めた多数の解析対象が不完全な検証のまま残ってしまった。

ほとんどの HERV は変異や欠失が多く、ウイルスとしての機能タンパク質をコードできる領域は残存していない。しかし、中には変異や欠失がほとんど見られず、ウイルス粒子を形成し得る配列を保持している HERV (HERVK の3ローカス) も見かける。これに対し、組織特異的プロモーター・エンハンサーとして働く HERV の LTR は少ない。HERV の存在の有無の個人差やエピジェネティックな制御能力の個人差が疾患感受性を規定する可能性が考えられる。今後は、更に充実することが予想されるヒトゲノムのアノテーションを参照しながら残りの解析対象の検討を進めると共に、Micro

Array の様に high- through-put な解析法の応用を検討してゆきたい。

今回の検討で明らかにできたのは、1p36.13 の HUERS-P3b および 5p15.33 の HERVK9 はヒトの脳で発現能を有する（少なくとも胎児脳では）ことであり、この二つは共にヒトとチンパンジーが分化する以前から霊長類ゲノムに進入していたと推定された。ニホンザルやアカゲザルが属する旧世界ザルでは検出できなかった事も合わせて考えると、進入時期は旧世界ザルが分岐した約 3000~2500 万年前から約 700~550 万年前までの間と考えられる。21q22.3 の HERVF(typeB)はヒトでのみ増幅されていることから約 700~550 万年前より新しく挿入されたコピーだと予想される。今後は long range PCR 法を用いた挿入部位の確認や、今回の増幅産物の塩基配列の比較も行うなど、更に詳細に検証する必要があると考えている。

6. 参考文献

- Conrad B, Weissmahr RN, Boni J, Arcari R, Schupbach J, Mach B. A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes. *Cell*. 1997 Jul 25;90(2):303-13.
- de Parseval N, Casella J, Gressin L, Heidmann T. Characterization of the three HERV-H proviruses with an open envelope reading frame encompassing the immunosuppressive domain and evolutionary history in primates. *Virology*. 2001 Jan 20;279(2):558-69.
- Karlsson H, Bachmann S, Schroder J, McArthur J, Torrey EF, Yolken RH. Retroviral RNA identified in the cerebrospinal fluids and brains of individuals with schizophrenia. *Proc Natl Acad Sci U S A*. 2001 Apr 10;98(8):4634-9.
- Komurian-Pradel F, Paranhos-Baccala G, Bedin F, Ounanian-Paraz A, Sodoyer M, Ott C, Rajoharison A, Garcia E, Mallet F, Mandrand B, Perron H. Molecular cloning and characterization of MSR-related sequences associated with retrovirus-like particles. *Virology*. 1999 Jul 20;260(1):1-9.
- Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, Keith JC Jr, McCoy JM. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature*. 2000 Feb 17;403(6771):785-9.
- Nakamura A, Okazaki Y, Sugimoto J, Oda T, Jinno Y. Human endogenous retroviruses with transcriptional potential in the brain. *J Hum Genet*. 2003;48(11):575-81. Epub 2003 Oct 15.
- Shiroma T, Sugimoto J, Oda T, Jinno Y, Kanaya F. Search for active endogenous retroviruses: identification and characterization of a HERV-E gene that is expressed in the pancreas and thyroid. *J Hum Genet*. 2001;46(11):619-25.

Sugimoto J, Matsuura N, Kinjo Y, Takasu N, Oda T, Jinno Y. Transcriptionally active HERV-K genes: identification, isolation, and chromosomal mapping. *Genomics*. 2001 Mar 1;72(2):137-44.

Tristem M. Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. *J Virol*. 2000 Apr;74(8):3715-30.