

組織特異的な遺伝子発現に影響を与えた転移因子の同定

島野桂伍¹ 里村和浩¹ 長田直樹¹ 遠藤俊徳^{1, a)}

概要: 本研究は進化的に組織特異的な遺伝子発現に影響を与えた転移因子の有無を明確にすることを目的とする。サル目 11 種、ツパイ目 1 種、ネズミ目 5 種、ウサギ目 2 種の哺乳綱計 19 種の、33 種の転移因子の位置情報データ計 71,260,578 個を材料として使用し、生物種間の転移因子の変異量、位置関係および全ゲノムにおける割合の比較を行った。その結果、ネズミ目において他種と変異量の平均が異なる転移因子が 6 種、生物特異的な転移因子が 12 種存在した。また、転移因子が遺伝子発現へ与えた影響を明らかにするため、ゲノムデータから遺伝子近傍を 5000bp ずつ区切りそれぞれの領域に存在する転移因子を同定し、発現プロファイルから遺伝子発現量が高い遺伝子近傍に有意に存在率が異なる転移因子を調べた。その結果、7 種の転移因子についてヒトの脳、肝臓、およびマウスの脾臓で遺伝子発現量が高い遺伝子近傍の転移因子の存在確率に有意な差が見られた。これらの遺伝子には転写調節、翻訳開始などに関わるものが多かった。

キーワード: 転移因子, 分子進化, 哺乳類

1. 背景・目的

転移因子は哺乳類のゲノムの 30~50%を構成しており、高い割合を占めている。転移因子は RNA 中間体を介して転移するレトロトランスポゾンと自身の配列を直接他の位置に転移させる DNA トランスポゾンで構成されている。[1]. 転移因子は、大多数が中立的な進化を示すため、宿主生物にとって有益な存在ではないと考えられてきた[2]. 近年の研究で、タンパク質コード領域や転写調節領域のような機能配列の一部が転移因子由来であるということが明らかになった[3]. また、哺乳類の乳腺や胎盤の形態形成進化に転移因子が関与したことが明らかにされた。[4,5]

一方でこのような機能的に貢献していることが明らかな転移因子はごく一部である。

本研究は哺乳類における組織特異的な遺伝子発現の変化に影響を与える転移因子を推定し、生物進化における影響を明らかにすることを目的とする。

2. 材料

RepeatMasker genomic datasets 上にある、種ごとにアノテーションされた転移因子のデータのうち、ヒト、ゴリラ、チンパンジー、オランウータン、テナガザル、ニホンザル、アカゲザル、コモンマーモセット、メガネザル、キツネザル、ガラゴ、ツパイ、リス、プレーリーハタネズミ、ラッ

ト、ハツカネズミ、モルモット、ウサギ、ナキウサギの合計 19 種の哺乳類の転移因子データを使用した。

また、遺伝子発現への影響を調べるため NCBI からヒト、ラット、マウスの GFF 形式の遺伝子データと脳、肝臓、脾臓の遺伝子の発現量を示したデータを使用した。

3. 結果・考察

3.1 各転移因子の特徴付け

生物種間の転移因子の変異量の平均値や転移因子の全ゲノムにおける割合を比較したところ、それぞれの転移因子は以下の特徴をもつものに分けられた。赤色の文字で記した転移因子は多峰性の分布を示した転移因子である。

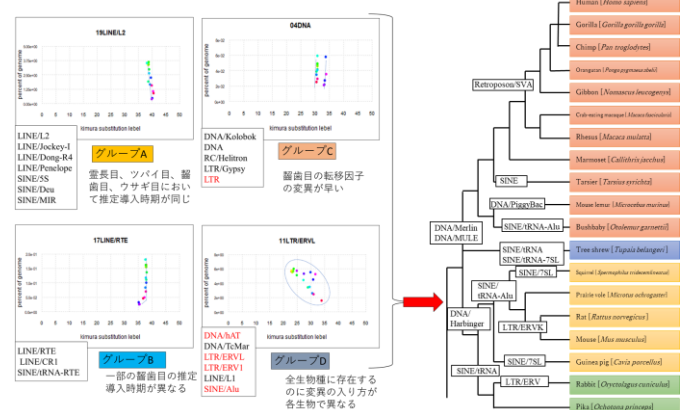


図 1 各転移因子を特徴ごとにまとめた図

各転移因子を特徴ごとにまとめると以下ようになった。
図 1 の x 軸の kimura substitution level は木村 2 パラメータ

¹ 北海道大学大学院情報科学院
a) endo@ist.hokudai.ac.jp

一モデル×100 で表されている。

- (A) 各生物種で転移因子の kimura substitution level(以下 ksl)の平均値がほぼ一致
- (B) ネズミ目のみ ksl の平均値が小さい
- (C) ネズミ目のみ ksl の平均値が大きい
- (D) 各生物種で ksl の平均値がバラバラである
- (E) 系統特異的に存在する

3.2 組織特異的な発現をする遺伝子近傍にある転移因子の同定

(1) Fisher の正確確率検定

組織特異的な遺伝子発現に影響を与えた転移因子を同定するため、ゲノムデータから遺伝子近傍を 5000bp ずつ区切りそれぞれの領域に存在する転移因子を同定し、R の Fisher の正確確率検定によって遺伝子発現量が高い遺伝子近傍に有意に存在率が異なる転移因子を求めた。

Human		脳		肝臓		脾臓	
転移因子名	グループ	-5000bp	5000bp	-5000bp	5000bp	-5000bp	5000bp
SINE/MIR	A	-4.25	-3.12	3.37	5.50	0.07	0.16
SINE/tRNA-RTE	B	0.38	1.14	-1.67	-3.79	0.45	-0.24
DNA/hAT	D	-0.61	-0.27	-2.82	-6.51	-0.24	0.58
DNA/TcMar	D	-0.16	5.98	-10.00	-12.00	0.23	0.07
LTR/ERV1	D	-3.83	-6.03	-1.25	-3.38	-0.62	-2.63
LINE/L1	D	-7.91	-7.29	-14.95	-16.71	-1.82	-2.44
SINE/Alu	D	12.53	14.95	6.66	3.45	7.30	7.25

図 2 Fisher の正確確率検定による遺伝子発現量が高い遺伝子近傍に有意に存在率が異なる転移因子の同定

上記の結果から、SINE/MIR, SINE/tRNA-RTE, DNA/hAT, DNA/TcMar, LTR/ERV1, LINE/L1, SINE/Alu の 7 種の転移因子が遺伝子発現の高い遺伝子近傍に対して有意に存在率が異なっていた。これらの遺伝子発現の高い遺伝子近傍に対して有意に存在率が異なる転移因子は近傍の遺伝子に対してエンハンサーまたはサイレンサーとして働くことで、発現の調節に関与した可能性がある。このうちの 5 つは特徴ごとに分けた際のグループ D に属している転移因子であった。グループ D は特定の生物種において機能を持ったことで変異が抑制された可能性がある。特に ksl のピークが 2 つある転移因子は遺伝子重複などによって新たな機能に関わっている可能性が高い。これと Fisher の正確確率検定の結果から、グループ D に属している転移因子は組織特異的な遺伝子発現の変化に対して関与した可能性が示唆された。

(2) 重回帰分析

Fisher の正確確率検定ではそれぞれの遺伝子近傍の転移因子数と遺伝子発現量の関係を 1 対 1 で見ていたが、転移因子全体の組織特異的な遺伝子発現に対する相互作用を調べるため、各遺伝子の Signal intensity の値を目的変数、各遺伝子近傍の転移因子数を説明変数として重回帰分析を行った。

Human		脳		肝臓		脾臓	
転移因子名	グループ	-5000bp	5000bp	-5000bp	5000bp	-5000bp	5000bp
SINE/MIR	A	8.96	2.42	11.15	9.25	0.53	0.28
SINE/tRNA-RTE	B	0.32	0.70	2.79	3.98	0.09	0.43
DNA/hAT	D	2.25	1.17	2.83	4.57	0.16	0.06
DNA/TcMar	D	0.38	2.71	9.66	11.30	0.00	0.00
LTR/ERV1	D	3.73	4.78	0.50	1.49	0.29	2.28
LINE/L1	D	20.74	20.24	9.14	11.77	6.00	4.40
SINE/Alu	D	39.94	37.97	17.77	12.81	20.60	15.40

図 3 重回帰分析による遺伝子発現量が高い遺伝子近傍に有意に存在率が異なる転移因子の同定

図の数値は検定におけるそれぞれの説明変数（ここでは遺伝子近傍の転移因子数）の p 値を対数で示したものであり、この数値が高いほど目的変数（ここでは遺伝子発現量）に対して影響を与えているということを示している。上記の結果から、Fisher の正確確率検定で有意だった 7 種の転移因子がそれぞれ遺伝子発現量と遺伝子近傍の転移因子数の間に相関があることがわかった。Fisher の正確確率検定と重回帰分析の結果から、転移因子全体の相互作用を考へてもこれらの転移因子は組織特異的な遺伝子発現量と遺伝子近傍の転移因子数の間に関係があると考えられる。

3.3 近傍に有意に存在率が異なる転移因子をもつ遺伝子の共通機能の同定

今までの Fisher の正確確率検定と重回帰分析において、遺伝子発現量と遺伝子近傍の転移因子数に相関がある転移因子の近傍の遺伝子群に共通する機能や局在が存在するかを調べるために DAVID version 6.8 によって gene ontology 解析を行った。

その結果、遺伝子近傍に SINE/tRNA-RTE, DNA/TcMar, LINE/L1 が有意に高確率で存在する遺伝子はそれぞれ、ヒトの肝臓、ヒトの肝臓、マウスの脾臓において組織特異的な機能を示した。また、これらの転移因子それぞれの近傍遺伝子群に共通して含まれている機能に、転写調節を担っているものや翻訳の調節を行うものがあつた。また、RPL11 という翻訳に関与する遺伝子において、近傍に存在する DNA/TcMar 配列内にこの遺伝子をターゲットとする転写因子結合部位である SP1 が 9 個あつた。これは DNA/TcMar が RPL11 の遺伝子発現の変化に関与していることを示唆している。これらから、各組織において特異的に発現する遺伝子の転写、翻訳を制御している遺伝子の発現の変化に対してこれらの転移因子が関与した可能性がある。

4. 結論

Fisher の正確確率検定で有意に遺伝子発現の高い遺伝子近傍に集まっている 7 種の転移因子は、重回帰分析および gene ontology 解析の結果から転写および翻訳の調節に関与している遺伝子に関与している可能性がある。これらから、今回同定した 7 種の転移因子は、各組織の形態形成におい

で特異的に働く遺伝子の発現の制御に関係した遺伝子に対して関与していたと考えられる。

参考文献

- [1] “Evolution and diversity of transposable elements in vertebrate genomes.” . <https://academic.oup.com/gbe/article/9/1/161/2965566>, (参照 2020-11-08).
- [2] “Genome-Wide Identification of Human Functional DNA Using a Neutral Indel Model” . <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.0020005>, (参照 2020-11-08).
- [3] “The impact of transposable elements in adaptive evolution “ . <https://onlinelibrary.wiley.com/doi/full/10.1111/mec.14794>, (参照 2020-11-08).
- [4] “Retrotransposons spread potential cis-regulatory elements during mammary gland evolution” . <https://academic.oup.com/nar/article/47/22/11551/5603223>, (参照 2020-5-15).
- [5] “Transposable element recruitments in the mammalian placenta: impacts and mechanisms” . <https://academic.oup.com/bfg/article/11/4/267/200284>, (参照 2020-11-08).