

令和 6 年 6 月 26 日現在

機関番号：14301

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K06119

研究課題名（和文）大規模並列レポーターアッセイ法による転移因子進化メカニズムの解明

研究課題名（英文）Elucidation of the mechanism of transposable element evolution using massively parallel reporter assay

研究代表者

井上 詞貴（Inoue, Fumitaka）

京都大学・高等研究院・特定准教授

研究者番号：60525369

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究では、転移因子機能の進化を明らかとするため、大規模並列レポーターアッセイ法(lentiMPRA)を用い、iPS細胞、神経前駆細胞における転移因子（MER11、MER34、MER52）のエンハンサー活性を大規模並列的に解析した。その結果、MER11の多くがiPS細胞において特異的に高いエンハンサー活性を持つことを見出した。さらに、得られたMER11バリエーションのエンハンサー活性は、MER11の進化系統関係と一致し、またエピジェネティック修飾（H3K27ac ChIP-seq, ATAC-seq）との相関から、霊長類進化過程におけるMER11エンハンサー活性の獲得・消失の変遷が明らかとなった。

研究成果の学術的意義や社会的意義

転移因子はヒトゲノムの半分近くを占め、近傍遺伝子の転写を制御するエンハンサーとして機能し、霊長類進化において重要な役割を果たすと考えられているが、その詳細は不明である。本研究により、比較的若い人内在性レトロウイルスMER11がiPS細胞で機能することを見出し、そのエンハンサー活性の大規模定量データを作成・公開することで、転移因子研究コミュニティへ貢献できる。また、転移因子のエンハンサー機能およびエピジェネティック機能を霊長類種間（ヒト、チンパンジー、マカクザル）で系統比較することにより、転移因子の機能性から見たゲノム進化の理解に繋がる。

研究成果の概要（英文）：To investigate the functional evolution of transposable elements, we analyzed the enhancer activity of endogenous retrovirus families (MER11, MER34, and MER52) in iPS cells and neural progenitor cells using a massively parallel reporter assay (lentiMPRA). We found that many of MER11 showed high enhancer activity specifically in iPS cells. Furthermore, the enhancer activities measured by MPRA and epigenetic activities (H3K27ac ChIP-seq, ATAC-seq) of MER11 evolutionary variants were consistent with their phylogenetic relationship. These results revealed the transition of gain and loss of MER11 enhancer activity during primate evolution.

研究分野：機能ゲノミクス

キーワード：エンハンサー MPRA 転移因子 エピゲノム ゲノム進化 iPS細胞 神経前駆細胞

1. 研究開始当初の背景

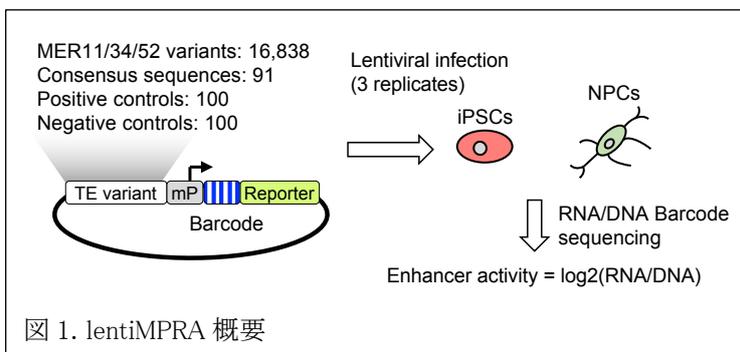
ヒトゲノムの 44%を占める転移因子は、機能を持たない”Junk DNA”と永らく考えられてきた。しかし、近年の大規模エピゲノム解析から、一部の転移因子ファミリーは生化学的機能(ヒストン修飾など)あるいは生物学的機能、例えば転写活性化を担うエンハンサー機能を持つことがわかってきた (Bourque et al., 2018; Kunarso et al., 2010)。転移因子はウイルス感染によりゲノムに組み込まれ、コピー&ペーストを繰り返して増殖し、中立的変異を蓄積してきた。しかしながら、このように生じた変異がいかんしてエンハンサー機能獲得に寄与し、進化過程で保存・変遷してきたのか、そしてそれがヒト進化にどう影響したのか、その詳細は不明である。これに対し、転移因子の配列比較解析や生化学的特徴の解析を行った研究はこれまでに多くあるものの、転移因子のエンハンサー活性を 1 塩基レベルで大規模機能解析した研究例は存在しない。申請者は先行研究において、ヒト ES 細胞から誘導した神経前駆細胞においてエピゲノムマップ (RNA-seq, ATAC-seq, ChIP-seq) を作成したところ (Inoue et al., 2019)、MER11、MER34、MER52 転移因子ファミリーがそれぞれ ES 細胞、神経前駆細胞特異的に活性型のエピジェネティックマーク(クロマチンアクセシビリティおよび H3K27ac 修飾)を獲得することを見出した(未発表)。これらのファミリーは真猿類(ヒト、チンパンジー、マカクを含む)進化の過程でゲノムに取り込まれ、コピーペーストを繰り返す過程で様々な変異を蓄積し、細胞特異的なエンハンサー機能を獲得してきたと考えられるが、その一つ一つの塩基置換が機能に及ぼす影響は依然として不明である。

2. 研究の目的

本研究では、申請者が開発してきた新規のレポーターアッセイ法である lentivirus-based Massively Parallel Reporter Assay (lentiMPRA) (Gordon et al., 2020)を用い、MER11、MER34、MER52 転移因子の全てのバリエーションのエンハンサー活性を定量比較することで、転移因子がどのようにしてエンハンサー機能を獲得し、進化的変遷を遂げたのかを、1 塩基レベルで明らかとする。

3. 研究の方法

まず、MER11、MER34、MER52 転移因子ファミリーの真猿類(ヒト、チンパンジー、マカク)ゲノム内バリエーション合計約 17,000 配列について、その DNA 断片を合成した (Twist Bioscience を利用)。加えて、ポジティブコントロールとして既知のエンハンサー配列、ネガティブコントロールとしてスクランブル化配列をそれぞれ 100 配列含めた。合成した DNA 断片を、プラスミドベクター内の最小プロモーター (mP) および GFP レポーター遺伝子上流に挿入した。同時に、15bp ランダム配列 (バーコードと呼ぶ) を GFP 遺伝子の 5' UTR へ挿入し、エンハンサー-バーコードライブラリを作成した (図1)。エンハンサーとバーコードの連鎖は、あらかじめ DNA-seq (NextSeq Mid-output PE150 を使用) により決定した。このライブラリをヒト iPSC 細胞 (iPSCs) にレンチウイルスを介して導入し、3 日間培養後、導入された DNA バーコードおよび転写された RNA バーコードを抽出し精製した。神経前駆細胞 (NPCs) を用いる場合は、dual-Smad 阻害法 (Chambers et al., 2009) により神経誘導を行い、72 時間後に DNA・RNA バーコードを抽出した。複製実験として、同様の操作を 3 回行った。DNA-seq、RNA-seq を行い (NextSeq High-output PE75 を使用)、各バリエーション毎のバーコード転写量 (DNA バーコード:RNA バーコード比) を測定することにより、バリエーションのエンハンサー活性を定量した。データプロセッシング (エラー、ノイズ除去、マッピング)、統計処理、定量解析は、バイオインフォマティクスツール MPRAflow (Gordon et al., 2020) および MPRAalyze (Ashuach et al., 2019) を用いて行った。得られたデータについて、各転移因子ファミリーの系統解析と照らして比較し、エンハンサー機能の変遷を検証した。



4. 研究成果

まず、各転移因子バリエーションのエンハンサー活性(連鎖する RNA バーコード:DNA バーコードカウント比)について、3 実験群間での再現性を検討したところ、高い相関を確認した(ピアソン相関関数=0.83, 図2A)。また、予想通り、既知エンハンサーポジティブコントロール(図2A, 緑)は高い活性を示し、スクランブル化配列ネガティブコントロール(図2A, 青)は低い活性を示したことから、lentiMPRA により得られた結果は信頼性の高いデータであると結論付けた。

次に、MER11、MER34、MER52 は、それぞれ MER11A, B, C、MER34A, A1, B, C, C2, C₃, D、

MER52A, C, D のサブファミリーとしてデータベース上にアノテーションされているため、各サブファミリーについて、iPSCs および NPCs におけるエンハンサー活性を検討したところ、MER11A, B, C (検証したバリエントのうちおよそ 50%) が iPSCs において特に高いエンハンサー活性を示すことが明らかとなった (図 2B)。したがって、MER11 は iPSCs において遺伝子の発現を制御するエンハンサーとして実際に機能する可能性が示唆された。

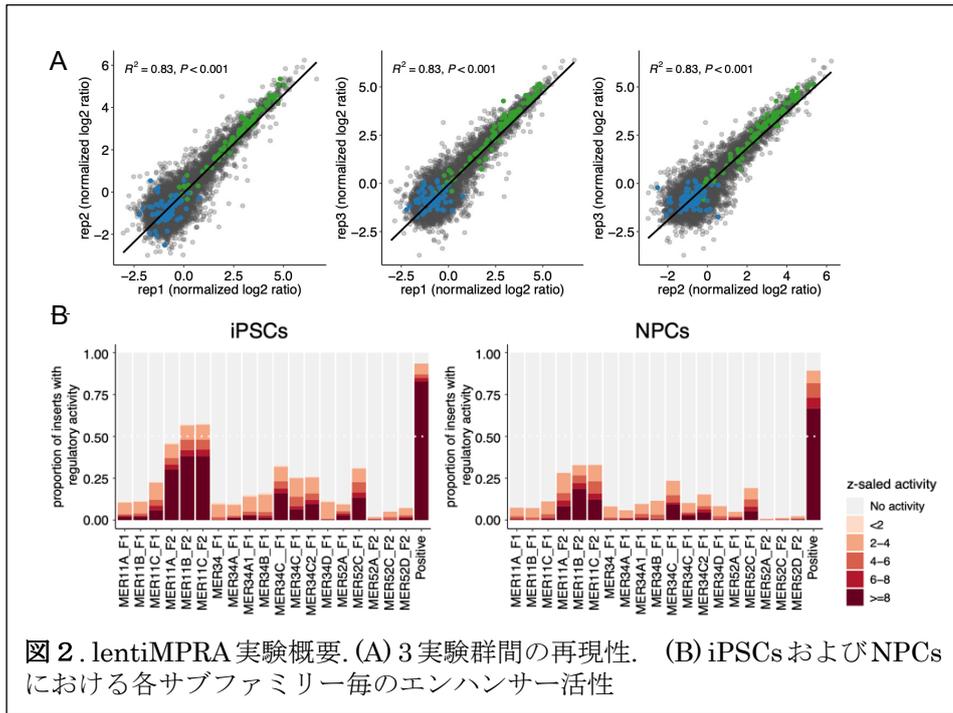


図 2. lentiMPRA 実験概要. (A) 3 実験群間の再現性. (B) iPSCs および NPCs における各サブファミリー毎のエンハンサー活性

そこで、以下 MER11 に注目することとした。MER11 は、既存の転移因子データベースにおいて MER11A/B/C として分類されているものの、全てのバリエントについて配列類似性を解析したところ、MER11A/B/C は実際には 66 個のクラスターに分割でき、MER11A/B/C 間で入り組んだ類似性を示すことがわかった (図 3A)。また、各クラスターについて系統解析を行ったところ、やはり相互に入り組んだ類似性を示した (図 3B)。このことから、既存の MER11A/B/C 分類は、その進化系統関係を正しく反映していないと結論付け、このクラスター間類似性から新たに 4 つのサブグループとして再分類することとした; MER11_G1 (図 3B 青)、MER11_G2 (図 3B 紫)、MER11_G3 (図 3B 緑)、MER11_G4 (図 3B 黄緑)。MER11_G1 は進化的に最も古く、また MER11_G4 は最も若いグループであると考えられる。

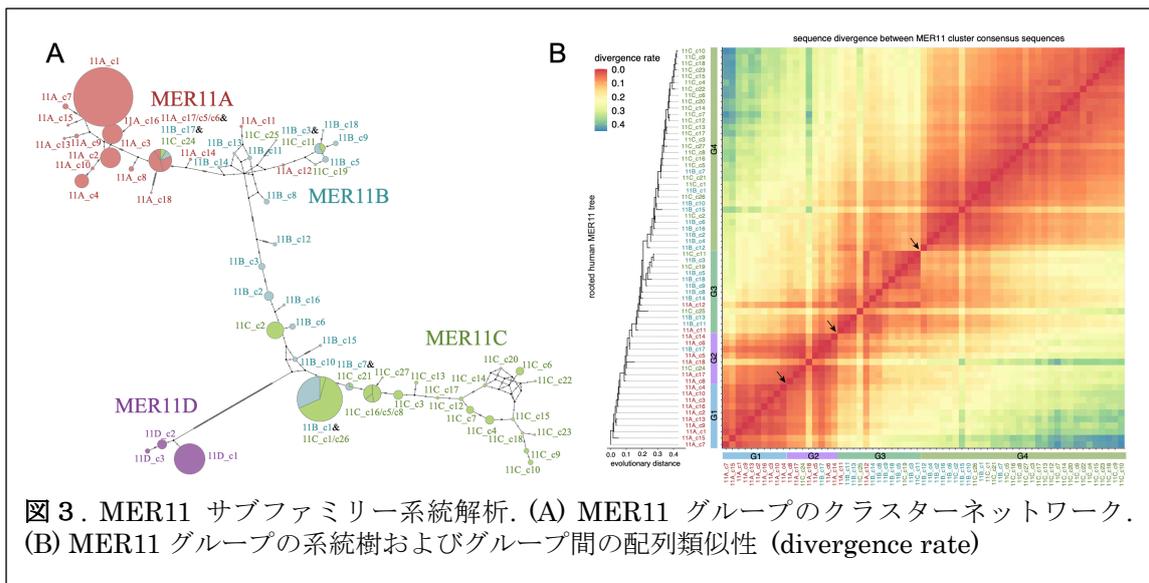
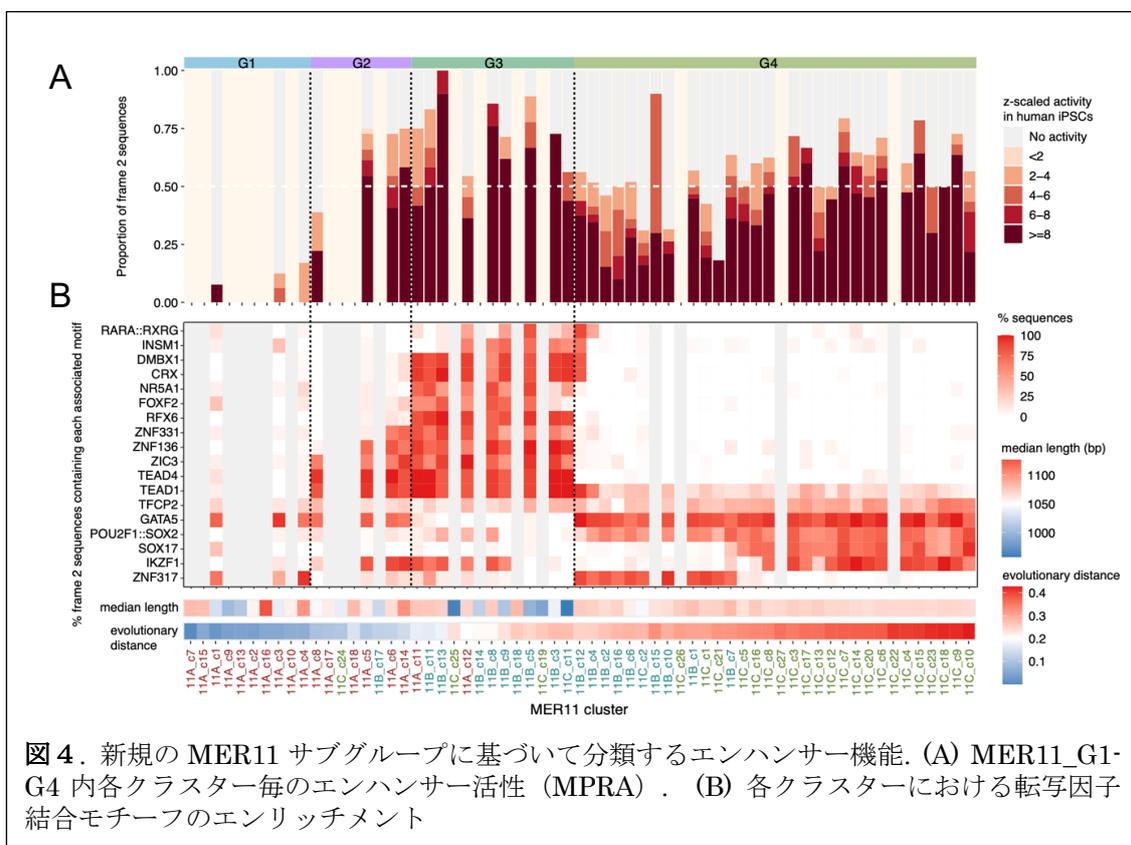


図 3. MER11 サブファミリー系統解析. (A) MER11 グループのクラスターネットワーク. (B) MER11 グループの系統樹およびグループ間の配列類似性 (divergence rate)

このようにして再定義した MER11_G1-4 に対し、lentiMPRA により測定したエンハンサー活性、エピジェネティック活性 (H3K27ac, アクセンビリティなど、data not shown)、および転写因子結合モチーフとの関連を検証した (図 4)。その結果、より若いグループである MER11_G3 および MER11_G4 が特に高いエンハンサー活性を持つことがわかった。また、MER11_G3 はホメオボックス転写因子 (DMBX1, CRX 等)、ZIC, TEAD 等のモチーフを持つ一方、MER11_G4 は POU, SOX 転写因子モチーフを持つことが

示された。すなわち、MER11_G3 はこれらの転写因子モチーフを維持する一方、一部の POU, SOX モチーフを持つバリエーションが MER11_G4 としてゲノム内に拡大し、その過程でホメオボックス, ZIC, TEAD モチーフを失ったと考えられる。このように、転写因子のゲノム内拡大に伴い、転写因子モチーフの獲得および消失が起こり、エンハンサー機能が変遷していったと考えられる。また、転写因子機能の変遷を正確に理解するためには、配列の系統関係に基づく正しい分類に基づいて、またエンハンサー活性定量データと合わせて検証する必要があることも示唆された。



引用文献

- Ashuach, T., Fischer, D.S., Kreimer, A., Ahituv, N., Theis, F.J., and Yosef, N. (2019). MPRAalyze: statistical framework for massively parallel reporter assays. *Genome Biol.* *20*, 183.
- Bourque, G., Burns, K.H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvák, Z., Levin, H.L., Macfarlan, T.S., et al. (2018). Ten things you should know about transposable elements. *Genome Biol.* *19*, 199.
- Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* *27*, 275-280.
- Gordon, M.G., Inoue, F., Martin, B., Schubach, M., Agarwal, V., Whalen, S., Feng, S., Zhao, J., Ashuach, T., Ziffra, R., et al. (2020). lentiMPRA and MPRAflow for high-throughput functional characterization of gene regulatory elements. *Nat. Protoc.* *15*, 2387-2412.
- Inoue, F., Kreimer, A., Ashuach, T., Ahituv, N., and Yosef, N. (2019). Identification and Massively Parallel Characterization of Regulatory Elements Driving Neural Induction. *Cell Stem Cell* *25*, 713-727.e10.
- Kunarso, G., Chia, N.Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y.S., Ng, H.H., and Bourque, G. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nat. Genet.* *42*, 631-634.

5. 主な発表論文等

〔雑誌論文〕 計10件（うち査読付論文 9件 / うち国際共著 10件 / うちオープンアクセス 9件）

1. 著者名 Georgakopoulos-Soares Ilias, Victorino Jesus, Parada Guillermo E., Agarwal Vikram, Zhao Jingjing, Wong Hei Yuen, Umar Mubarak Ishaq, Elor Orry, Muhwezi Allan, An Joon-Yong, Sanders Stephan J., Kwok Chun Kit, Inoue Fumitaka, Hemberg Martin, Ahituv Nadav	4. 巻 2
2. 論文標題 High-throughput characterization of the role of non-B DNA motifs on promoter function	5. 発行年 2022年
3. 雑誌名 Cell Genomics	6. 最初と最後の頁 100111 ~ 100111
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.xgen.2022.100111	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 該当する
1. 著者名 Kreimer Anat, Ashuach Tal, Inoue Fumitaka, Khodaverdian Alex, Deng Chengyu, Yosef Nir, Ahituv Nadav	4. 巻 13
2. 論文標題 Massively parallel reporter perturbation assays uncover temporal regulatory architecture during neural differentiation	5. 発行年 2022年
3. 雑誌名 Nature Communications	6. 最初と最後の頁 1504
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41467-022-28659-0	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 該当する
1. 著者名 Koesterich Justin, An Joon-Yong, Inoue Fumitaka, Sohota Ajuni, Ahituv Nadav, Sanders Stephan J., Kreimer Anat	4. 巻 24
2. 論文標題 Characterization of De Novo Promoter Variants in Autism Spectrum Disorder with Massively Parallel Reporter Assays	5. 発行年 2023年
3. 雑誌名 International Journal of Molecular Sciences	6. 最初と最後の頁 3509 ~ 3509
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/ijms24043509	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 該当する
1. 著者名 Whalen Sean, Inoue Fumitaka, Ryu Hane, et. al.	4. 巻 111
2. 論文標題 Machine learning dissection of human accelerated regions in primate neurodevelopment	5. 発行年 2023年
3. 雑誌名 Neuron	6. 最初と最後の頁 857 ~ 873.e8
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.neuron.2022.12.026	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 該当する

1. 著者名 Ilias Georgakopoulos-Soares, Chengyu Deng, Vikram Agarwal, Candace S. Y. Chan, Jingjing Zhao, Fumitaka Inoue & Nadav Ahituv	4. 巻 14
2. 論文標題 Transcription factor binding site orientation and order are major drivers of gene regulatory activity	5. 発行年 2023年
3. 雑誌名 Nature Communications	6. 最初と最後の頁 2333
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41467-023-37960-5	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Keough Kathleen C., Whalen Sean, Inoue Fumitaka, et al.	4. 巻 380
2. 論文標題 Three-dimensional genome rewiring in loci with human accelerated regions	5. 発行年 2023年
3. 雑誌名 Science	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1126/science.abm1696	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Liu Jiayi, Ashuach Tal, Inoue Fumitaka, Ahituv Nadav, Yosef Nir, Kreimer Anat	4. 巻 52
2. 論文標題 Optimizing sequence design strategies for perturbation MPRA: a computational evaluation framework	5. 発行年 2024年
3. 雑誌名 Nucleic Acids Research	6. 最初と最後の頁 1613 ~ 1627
掲載論文のDOI (デジタルオブジェクト識別子) 10.1093/nar/gkae012	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Capauto Davide, Wang Yifan, Wu Feinan, Norton Scott, Mariani Jessica, Inoue Fumitaka, Crawford Gregory E., Ahituv Nadav, Abyzov Alexej, Vaccarino Flora M.	4. 巻 14
2. 論文標題 Characterization of enhancer activity in early human neurodevelopment using Massively Parallel Reporter Assay (MPRA) and forebrain organoids	5. 発行年 2024年
3. 雑誌名 Scientific Reports	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41598-024-54302-7	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Feng Yuanqing, Xie Ning, Inoue Fumitaka, Fan Shaohua, Saskin Joshua, Zhang Chao, Zhang Fang, Hansen Matthew E. B., Nyambo Thomas, Mpoloka Sununguko Wata, Mokone Gaonyadiwe George, Fokunang Charles, Belay Gurja, Njamshi Alfred K., Marks Michael S., Oancea Elena, Ahituv Nadav, Tishkoff Sarah A.	4. 巻 56
2. 論文標題 Integrative functional genomic analyses identify genetic variants influencing skin pigmentation in Africans	5. 発行年 2024年
3. 雑誌名 Nature Genetics	6. 最初と最後の頁 258 ~ 272
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41588-023-01626-1	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Chen Xun, Zhang Zicong, Yan Yizhi, Goubert Clement, Bourque Guillaume, Inoue Fumitaka	4. 巻 -
2. 論文標題 Cryptic endogenous retrovirus subfamilies in the primate lineage	5. 発行年 2023年
3. 雑誌名 BioRxiv	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1101/2023.12.07.570592	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

〔学会発表〕 計4件 (うち招待講演 4件 / うち国際学会 0件)

1. 発表者名 井上詞貴
2. 発表標題 転移因子エンハンサーバリエーションの大規模機能解析
3. 学会等名 第45回日本分子生物学会年会 (招待講演)
4. 発表年 2022年

1. 発表者名 井上詞貴
2. 発表標題 大規模並列レポーターアッセイによるノンコーディングゲノム解読
3. 学会等名 NGS EXPO (招待講演)
4. 発表年 2022年

1. 発表者名 井上詞貴
2. 発表標題 Massively parallel dissection of regulatory code underlying neural differentiation
3. 学会等名 第56回日本発生生物学会（招待講演）
4. 発表年 2023年

1. 発表者名 井上詞貴
2. 発表標題 大規模並列レポーターアッセイ(MPRA)による転移因子エンハンサーの機能解析
3. 学会等名 第6回転移因子研究会（招待講演）
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>京都大学ASHBi 井上研究室（Bourqueグループ） https://ashbi.kyoto-u.ac.jp/lab-sites/inoue_lab/</p>
--

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	BOURQUE GUILLAUME (Bourque Guillaume) (80890566)	京都大学・高等研究院・主任研究者 (14301)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	C H E N X U N (Chen Xun) (30885158)	京都大学・高等研究院・特定助教 (14301)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
カナダ	マギル大学			