

## 「統合トランスポゾンデータベースを活用した癌トランスクリプトーム」

### 【研究概要】

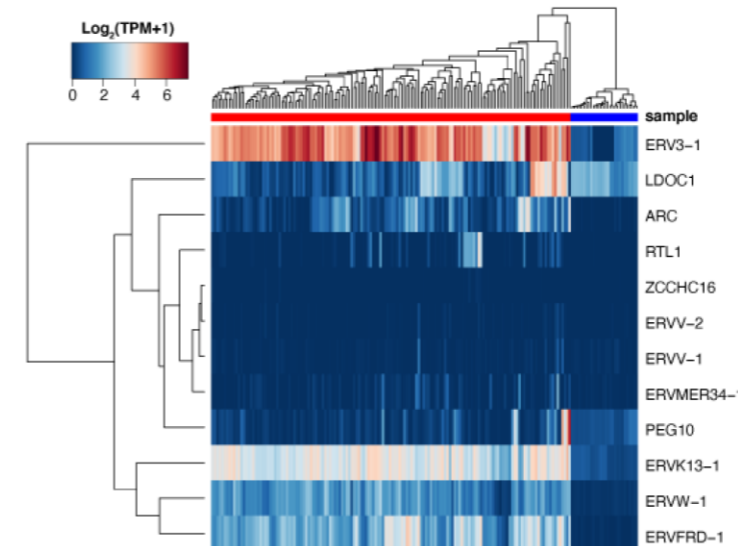
ヒトを含む哺乳類ではトランスポゾンに由来する配列が全ゲノムのおよそ半分を占めている。トランスポゾンは宿主ゲノムに寄生するジャンクDNAだと考えられてきたが、近年のシーケンス技術の進展により、トランスポゾン中にさまざまな機能を担うエレメントが存在し、発癌、神経性疾患、ウイルス感染防御などに関与する可能性があることが明らかとなった。しかしゲノムに存在するトランスポゾンの多くは、アノテーションがなく、その機能もよくわかっていない。我々はこれまでに、トランスポゾン的一种であるLTRトランスポゾンについて、19種の哺乳類ゲノムを対象に網羅的に同定・アノテーションしたデータベース (gEVE, <http://geve.med.u-tokai.ac.jp>) を構築し、宿主で獲得した新しい機能を明らかにしてきた (Nakagawa et al, Database 2016)。本研究ではgEVEデータベースを活用し、米国がんゲノムデータベースThe Cancer Genome Atlas (TCGA) に集められている造血器腫瘍患者由来の大規模転写配列データ (RNA-seq) を合わせて解析する。そして、造血器腫瘍の病態の進展に関与するトランスポゾンの同定を行う。そのようなトランスポゾンを制御することで、造血器腫瘍の予後の予測、そして治療にも貢献することを目指す。

### 【研究成果】

米国がんゲノムデータベースTCGAから、急性骨髄性白血病 (AML) 患者由来のRNA-seqサンプルと患者情報を入手した。ヒトゲノムの遺伝子アノテーションとgEVEデータベースによるEVEアノテーション情報を用いて、各AML患者のウイルス様配列の発現解析を行った。上記の中で予後不良・良好の指標となるデータがないかを解析した。AML患者由来のサンプルを解析した結果、レトロウイルス由来の配列のなかで、とくにERV3-1という膜タンパク質に由来する遺伝子が多くのAML患者で発現向上していることがわかった (Nakagawa et al. Gene, 2021)。実際に抗体を使ってERV3-1がタンパク質として発現していることも明らかにした。一方で、ERV3-1の白血球での機能や、その発現レベルと予後の関係は明らかにならなかった。また、本解析の結果、ERV3-1以外にも一部のAML患者では高発現しているHERV由来の配列があることが分かり、現在研究を継続している。加えて、ゲノムに存在するウイルス様の配列に関する機能・進化についてさまざまな研究を進めている。ヒトでは見つからないが、チンパンジーなどの類人猿に見つかったレトロウイルス膜タンパク質由来の配列がウイルス感染を防ぐ役割を担っていることを発見した (Miyake et al. PNAS 2022)。レトロウイルスの逆転写酵素に由来する配列が単孔類2種の間で蛋白質コード領域が強く保存されていることがわかり、精巢特異的に発現していることから、遺伝子としてなんらかの機能を担っていることを報告した (Kitao et al. Virus Evol 2022)。また、ヒトなどの胎盤形成に重要な機能を果たすと考えられているレトロウイルス膜タンパク質由来のsyncytin-2遺伝子が、新世界ザルではあまり保存されておらず、また融合活性も低下していることを示し、別な遺伝子が胎盤発生に寄与している可能性を示した (Shoji et al. FEBS Open Bio 2023)。このような内在性レトロウイルス由来の遺伝子と胎盤進化に関する総説も報告した (Imakawa et al. Cells 2022)。以上の研究に加えて、新型コロナウイルスに関係する研究も進めていて、特に塩基変異とそれに伴うウイルスの性状の変化やゲノムデータの効果的な保存方法について報告した (Takada et al. iScience 2023; Kimura et al. iScience 2023; Kryukov et al. Patterns 2022)。

### 【今後の展望】

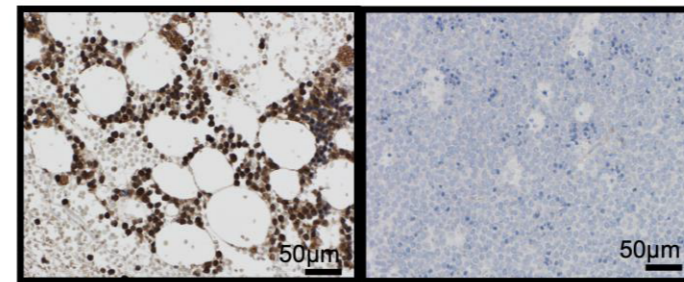
いままでの解析をベースにして、造血器腫瘍患者で発現しているレトロウイルス由来の配列について、一部の患者でのみ発現している配列について、更に解析を進め、予後との関連や、新たな治療法への活用などについて様々検討を行う予定である。加えて、造血器腫瘍以外の癌、もしくはその他の疾患についても、内在性レトロウイルスの発現との関係について明らかにする計画である。



AML 患者（赤）と健常者（青）の白血球由来の RNA-seq データ解析



プロジェクトリーダー：中川 草  
*Nakagawa So*  
医学部医学科基礎医学系分子生命科学 准教授



ERV3-1 のタンパク質発現を免疫染色法により解析 (左 positive, 右 negative)

### Selected Papers

- Shoji H, Kitao K, Miyazawa T, Nakagawa S. Potentially reduced fusogenicity of syncytin-2 in New World monkeys. *FEBS Open Bio*, 13(3): 459-467 (2023).
- Takada K, Ueda MT, Shichinohe S, Kida Y, Ono C, Matsuura Y, Watanabe T, Nakagawa S. Genomic diversity of SARS-CoV-2 can be accelerated by mutations in the nsp14 gene. *iScience*, 26(3): 106210 (2023).
- Ohkura S, Shimizu M, Horie M, Nakagawa S, Osanai H, Miyagawa Y, Morita R. Characterization of megabat-favored, CA-dependent susceptibility to retrovirus infection. *Journal of Virology*, 97(3): e0180322 (2023).
- Kimura I, Yamasoba D, Nasser H, Zahradnik J, Kosugi Y, Wu J, Nagata K, Uriu K, Tanaka YL, Ito J, Shimizu R, Tan TS, Butlertanaka EP, Asakura H, Sadamasu K, Yoshimura K, Ueno K, Takaori-Kondo A, Schreiber G, The Genotype to Phenotype Japan (G2P-Japan) Consortium, Toyoda M, Shirakawa K, Irie T, Saito A, Nakagawa S, Ikeda T, Sato K. The SARS-CoV-2 spike S375F mutation characterizes the Omicron BA.1 variant. *iScience* 25: 105720 (2022).
- Kitao K, Miyazawa T, Nakagawa S. Monotreme-specific conserved proteins derived from retroviral reverse transcriptase. *Virus Evolution* 8(2): veac084 (2022).
- Sakaguchi S, Urayama S, Takaki Y, Hirotsuna K, Wu H, Suzuki Y, Nunoura T, Nakano T, Nakagawa S. NeoRdRp: A comprehensive dataset for identifying RNA-dependent RNA polymerase of various RNA viruses from metatranscriptomic data. *Microbes and Environments* 37(2): ME22001 (2022).
- Miyake A, Ngo MH, Wulandari S, Shimojima M, Nakagawa S, Kawasaki J, Nishigaki K. Convergent evolution of antiviral machinery derived from endogenous retrovirus truncated envelope genes in multiple species. *Proceedings of the National Academy of Sciences of the United States of America* 119(26): e2114441119 (2022).
- Takeuchi K, Ikeda Y, Senda M, Harada A, Okuwaki K, Fukuzawa K, Nakagawa S, Yu HY, Nagase L, Imai M, Sasaki M, Lo YH, Ito D, Osaka N, Fujii Y, Sasaki AT, Senda T. The GTP-responsiveness of PI5P4K  $\beta$  is evolved by a compromised trade-off between activity and specificity. *Structure* 30(6): 886-899 (2022).
- Kryukov K, Jin L, Nakagawa S. Efficient compression of SARS-CoV-2 genome data using Nucleotide Archival Format (NAF). *Patterns* 3(9): 100562 (2022).
- Imakawa K, Kusama K, Kaneko-Ishino T, Nakagawa S, Kitao K, Miyazawa T, Ishino F. Endogenous Retroviruses and Placental Evolution, Development, and Diversity. *Cells* 11(15): 2458 (2022).
- Kryukov K, Imanishi T, Nakagawa S. Nanopore sequencing data analysis of 16S rRNA genes using GenomeSync-GSTK system. *Methods in Molecular Biology*, 2632: 215-226 (2023).