

令和 4 年 6 月 23 日現在

機関番号：12601

研究種目：若手研究

研究期間：2019～2021

課題番号：19K16740

研究課題名（和文）血球分化の各階層のクロマチン構造データを用いた成人T細胞白血病発症機構の解明

研究課題名（英文）Elucidation of the pathogenesis of adult T-cell leukemia using chromatin structural data at various levels of hematopoietic differentiation

研究代表者

田中 梓（Tanaka, Azusa）

東京大学・大学院医学系研究科（医学部）・特別研究員

研究者番号：70749796

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）：成人T細胞白血病（ATL）は、ヒトT細胞白血病ウイルス（HTLV-1）の感染により発症する難治性のT細胞悪性腫瘍である。本研究ではATL症例細胞のクロマチン構造を詳細に解析することで、HTLV-1感染細胞がATLを発症するメカニズムの解明を目指した。課題遂行のためには健常細胞との比較解析が必須であり、まずサンプル間の比較解析の手法を開発した。またその手法をATL細胞の解析に用いるとATLはT細胞白血病であるにも関わらず、骨髄球細胞に類似したクロマチン構造を有する症例があることが明らかとなり、「骨髄球系細胞に近い形質を有するT細胞白血病細胞」の存在の可能性が示唆された。

研究成果の学術的意義や社会的意義

HTLV-1は母子感染するレトロウイルスで、数千年前から人類と共存してきた。感染者の90%は無症候のまま健康状態を保っているが、感染者の5%は成人T細胞白血病（ATL）として知られる難治性の白血病やリンパ腫を発症する。HTLV-1は造血幹細胞を含め様々な細胞に感染し、さらに感染を維持したまま分化することが報告されているが、ATL発症メカニズムの関しては未だ不明な点も多い。本研究課題では健常人由来の細胞とのクロマチン構造の比較解析により、T細胞以外の要素を有するATL症例の存在が示唆された。

研究成果の概要（英文）：Human T-cell leukemia virus type 1 (HTLV-1) causes adult T-cell leukemia-lymphoma (ATL). The leukemic cells typically have a CD4+ memory T cell phenotype, but the mechanisms underlying this process are not presently understood. To clarify the mechanism of leukemia development, we focused on the chromatin structure, which is a central mechanism of cell regulation, and investigated the characteristics of ATL cells. We first developed a method for comparative analysis between samples. Next, we analyzed the chromatin accessibility landscape of HTLV-1-infected cells from ATL cases by comparing the ATAC-seq data with those from 13 types of hematopoietic cells from healthy donors using this method. It was found that in some cases, although ATL is a T-cell leukemia, the chromatin structure resembled to that of myeloid cells, suggesting the possibility of the existence of "T-cell leukemia cells with traits similar to those of myeloid cells".

研究分野：エピゲノム

キーワード：成人T細胞白血病ウイルス エピゲノム

## 1. 研究開始当初の背景

成人 T 細胞白血病 (ATL) 細胞はその 90%以上が CD4 陽性メモリー T 細胞である (Richardson et al., J. Virol., 1990)。しかしながら、ATL の病因ウイルスである成人 T 細胞白血病ウイルス (HTLV-1) は造血幹細胞 (HSC) を含め様々な細胞に感染し、さらに感染を維持したまま分化することが報告されている (Furuta et al., Plos Pathog., 2018)。HTLV-1 にコードされる *HTLV-1 bZIP factor (HBZ)* や *tax* という遺伝子の機能解析はこれまで多数報告されている (Tanaka and Matsuoka, Front Microbiol, 2018)。しかし、最も基本的情報であるはずの ATL 細胞の由来に関しては見解が分かれ、ATL 細胞の由来が CD4 陽性メモリー T 細胞なのか、もしくは分化の上流にいる細胞になんらかの異常が起こり CD4 陽性メモリー T 細胞という特定の分画へ分化してしまったのか、明らかになっていない。また ATL は HTLV-1 感染から 60 年という長期潜伏期間を経て一部の感染者のみに発症するため、実験的に感染～白血病発症までの過程を追うことは不可能である。HTLV-1 のウイルスタンパク質の機能解析を目的として作成されたトランスジェニックマウスは、そのほとんどが T 細胞で標的のウイルスタンパク質を発現するように改変されたものである。しかしながら ATL 細胞の由来が明確に同定されていないため、本来のウイルスタンパク質の機能が完全に理解できていない可能性がある。このような背景から、ATL 細胞の由来を同定することが求められてきた。

## 2. 研究の目的

成人 T 細胞白血病 (ATL: Adult T-cell Leukemia) は、ヒト T 細胞白血病ウイルス (HTLV-1: Human T-cell Leukemia Virus type 1) の感染により発症する難治性の T 細胞悪性腫瘍である。ATL 細胞の由来を同定することは、治療標的細胞を明確にするのみならず、ATL 細胞と比較すべき細胞分画がより正確になることで、ATL 発症メカニズムを高い精度で解明することができる。しかし未だ ATL 細胞の由来がどの細胞分画であるか明確になっていない。細胞の過去の状態、現在、未来の運命決定を説明するには遺伝子発現だけでなく、エピゲノムの情報が必須である。そこで、ATL のクロマチンデータとデータベース上に蓄積された様々な血球系のクロマチンデータを詳細に比較解析することで、ATL の由来となった血球分画がどのような異常を経て ATL 発症へと至るのかを解明することを目的に研究を進めた。

## 3. 研究の方法

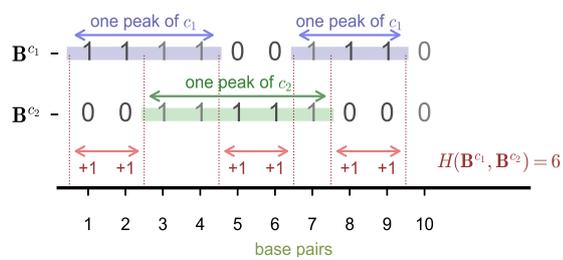
本研究では正常細胞と ATL 細胞の比較解析にオープンクロマチンの情報を利用することとし、ATAC-seq のデータを利用した。次世代シーケンサーから出力されるオープンクロマチンのデータをどう処理・解析すれば複数サンプル間の比較解析をより正確に・容易に行うことができるかを第一の課題として本研究を進めた。

## 4. 研究成果

(1) サンプル間距離の計算にハミング距離を導入することで比較解析を容易にした

ATAC-seq 法は、全ゲノム領域のクロマチン構造を網羅的に解析する手法の 1 つであるが、複数サンプルのデータ比較解析の手法は論文により様々であった。ATAC-seq 法では、Tn5 という酵素により、オープンクロマチン領域特異的に DNA が切断される。そのため、切断された DNA が多いということは即ち、クロマチンがオープンであることを意味する。このような断片化された DNA

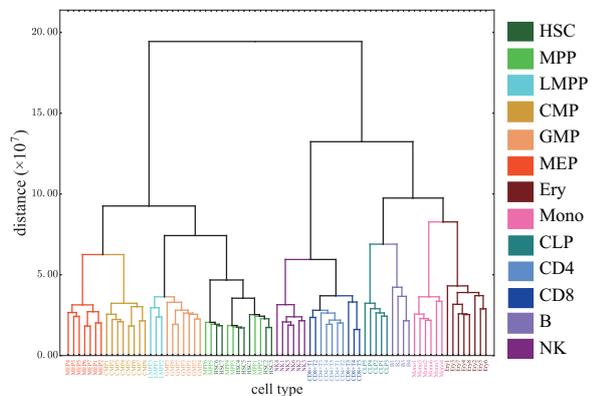
が集積している領域は「ピーク」と呼ばれ、ピークを自動検出するための様々なツールが開発されている。本研究では、全ゲノム領域を「ピーク」と被るか否か、つまり、「オープンクロマチン領域にあるか否か」で、0, 1 で表記し、2 サンプル間の距離をこの 0, 1 表記の違いの総和で表すハミング距離という概念を導入した (図 1)。



【図 1】 Hamming 距離の概念図

### (2) 距離計算に用いるピーク領域選択のためのオプションの最適化

ピークの数、ピーク抽出の際にどのようなパラメータを用いたか、また、ノイズとも思われるような小さなピークも、ピークとして解析に用いるのかに大きく依存する。そのため、(1) で述べたハミング距離を適用するためには、これらの条件の良し悪しを詳細に検討する必要がある。そこで、健康人由来の造血幹細胞 (HSC)、多能性造血前駆細胞 (MPP)、リンパ球系多能性前駆細胞 (LMPP)、骨髄球系共通前駆細胞 (CMP)、顆粒球・マクロファージ前駆細胞 (GMP)、リンパ球系共通前駆細胞 (CLP)、CD4 陽性 T 細胞、CD8 陽性 T 細胞、B 細胞、ナチュラルキラー細胞、巨核球・赤芽球前駆細胞 (MEP)、単球 (Mono)、赤血球 (Ery) の 13 種類 77 サンプルの血球分画の ATAC-seq データを用いて、どのような設定にすると、最も細胞の状態を正確に反映するのか独自の評価指標を用いて検討した。その結果、特定のピークコールの条件 (MACS2 アルゴリズムを  $p=0.01$  という条件で用いる) でピークコールを行い、その結果出力される数万の“ピークとして認識された領域”を信頼度順に並べ、その上位 64000 までのピークを用いることが最適という結論が得られた。



【図 2】 クラスタリング結果の一例

次に、上述の条件でサンプル間のハミング距離を計算し、ward 法を用いてクラスタリング解析を行なった。図 2 のように、HSC と MPP の 2 つは先行研究同様分類が完璧ではないが、それ以外の細胞種類に関しては、細胞の状態を正確に反映する分類結果となり、この手法の有用性が示された。

### (3) データ量変動の影響を検証

次に開発した手法がデータ量の変化にどの程度影響を受けるのかを調べた。その結果、シーケンサーから得られたデータ量が半分になっても、クラスタリング解析の結果に大きな影響は及ぼさないことがわかり、データ量の変動にも強いことがわかった。また既存の手法と本手法の計算量の比較も行い、10 倍以上計算時間が短いことも明らかとなった。

### (4) 白血病細胞を用いた解析

最後に、多数の先行研究がある急性骨髄性白血病 (AML)、慢性リンパ性白血病 (CLL) のデータを用いて、これらの白血病細胞がどの細胞種に類似しているのかをハミング距離のデータを元に同定した。その結果、AML のデータは先行研究での結果とよく一致し、B 細胞由来とされている CLL 症例データはハミング距離から推察した結果、B 細胞に最も近いという結果が得られた。ATL 細胞でも同様の解析を行なった結果、ATL は T 細胞へと分化した形跡があるにも関わらず、単球分画に非常に近いクロマチン構造を有する症例があることが明らかとなった。そこで改めて ATL 細胞の遺伝子発現データの再解析を行ったところ、一部の ATL 細胞では単球のマーカーとして用いられている遺伝子 CD14 の発現が高く、そのほか骨髄球系 (T 細胞はリンパ球系の

細胞であり単球は骨髄球系に分類される) の細胞に高発現している CD11b の高発現も確認され、「単球などの骨髄球系細胞に近い形質を有する T 細胞白血病細胞」の存在の可能性が示唆された。

本研究で作成した解析パイプラインは <https://github.com/tanakanishi/findclosest> で公開し、研究成果は論文として報告した (PLoS Comput Biol. 2020 Nov 30;16(11):e1008422.)。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 Tanaka Azusa, Ishitsuka Yasuhiro, Ohta Hiroki, Fujimoto Akihiro, Yasunaga Jun-ichirou, Matsuoka Masao	4. 巻 16
2. 論文標題 Systematic clustering algorithm for chromatin accessibility data and its application to hematopoietic cells	5. 発行年 2020年
3. 雑誌名 PLOS Computational Biology	6. 最初と最後の頁 e1008422
掲載論文のDOI（デジタルオブジェクト識別子） 10.1371/journal.pcbi.1008422	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 -

1. 著者名 Jing Hao Wong, Daichi Shigemizu, Yukiko Yoshii, Shintaro Akiyama, Azusa Tanaka, Hidewaki Nakagawa, Shu Narumiya & Akihiro Fujimoto	4. 巻 11
2. 論文標題 Identification of intermediate-sized deletions and inference of their impact on gene expression in a human population	5. 発行年 2019年
3. 雑誌名 Genome Medicine	6. 最初と最後の頁 1-16
掲載論文のDOI（デジタルオブジェクト識別子） 10.1186/s13073-019-0656-4	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 -

〔学会発表〕 計1件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 田中梓、安永純一郎、藤本明洋、松岡雅雄
2. 発表標題 成人T細胞白血病細胞のクロマチン構造解析
3. 学会等名 第5回日本HTLV-1学会学術集会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
---------------------------	-----------------------	----

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
---------	---------