

令和 3 年 6 月 21 日現在

機関番号：12102

研究種目：研究活動スタート支援

研究期間：2019～2020

課題番号：19K24361

研究課題名(和文) 集団ゲノムデータに基づき変異を避けることによる遺伝子治療用gRNA設計法の確立

研究課題名(英文) Establishment of gRNA design method for gene therapy by avoiding mutations based on population genome data

研究代表者

土屋 貴穂 (Tsuchiya, Takaho)

筑波大学・医学医療系・助教

研究者番号：70853167

交付決定額(研究期間全体)：(直接経費) 2,200,000円

研究成果の概要(和文)：転写因子認識配列の多様性を捉えることは、複雑な遺伝子発現制御機構を紐解く上で重要である。本研究課題では、従来手法では捉えられなかった転写因子認識配列の多様性を捉えることを目指し、ChIP-Atlasに集約された約10万のChIP-seqデータに対して我々が開発したMOCCSを適用した。我々のアプローチにより、様々な細胞型における各転写因子の認識配列とその結合特異性スコアを解析し、細胞型ごとの転写因子認識配列の多様性が明らかとなった。

研究成果の学術的意義や社会的意義

本研究課題では、様々な細胞型における各転写因子の認識配列とその結合特異性スコアを解析し、細胞型ごとの転写因子認識配列の多様性が明らかとなった。本研究結果で得られた知見は、転写因子を標的とした創薬など応用面にも貢献し得る。そして、本研究結果で得られた知見を社会発信するために、現在、生物種や細胞型による転写因子認識配列の多様性を閲覧・比較できるデータベースを構築している。論文投稿の準備も行っている。本研究課題の期間終了後も、種間・細胞型間の転写因子認識配列の比較と転写調節共益因子の同定や変異の影響予測などの応用に発展させることを目指す。

研究成果の概要(英文)：The diversity of recognition sequences in transcription factors is important for gene expression regulation. In this project, we applied our MOCCS to about 100,000 ChIP-seq data collected in ChIP-Atlas in order to capture the diversity of transcription factor recognition sequences that could not be captured by conventional methods. Using our approach, we analyzed the recognition sequences of each transcription factor and their binding specificity scores in various cell types, and revealed the diversity of transcription factor recognition sequences in each cell type.

研究分野：バイオインフォマティクス

キーワード：バイオインフォマティクス エピジェネティクス データベース 転写因子認識配列 細胞型 一塩基多型

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

## 1. 研究開始当初の背景

近年、CRISPR/Cas9 guide RNA (gRNA) を用いたゲノム編集技術は細胞工学だけでなく遺伝子治療の手段としても応用され始めている (Cathomen et al., 2018)。遺伝子治療では gRNA がゲノム上の相補的な標的配列にハイブリダイズして疾患原因遺伝子の発現低下などを引き起こし薬効が発揮される。薬効を発揮する gRNA の配列設計は情報科学の課題として数多く研究されてきており、リファレンスゲノムに基づいた配列相同性検索によりオフターゲット結合を除く目的で CRISPRdirect などが開発されている (Naito et al., 2014)。しかしながら、従来のアルゴリズムでは変異などゲノム配列の個人差を考慮していないという問題があった。単一のリファレンスゲノムに基づいた設計だけでは、リファレンスゲノムと異なる配列を持っている人にはハイブリダイズしないので治療失敗につながる。特に人類集団レベルで差が顕著な場合には対象患者層を大きく狭める危険性もある。そのため、ゲノム配列の個人差を考慮した gRNA 配列設計が必要とされていた (Lessard et al., 2017)。

## 2. 研究の目的

より広い患者層で薬効を発揮することのできる gRNA 設計手法を開発すること。

## 3. 研究の方法

データベースに集積されている集団ゲノムデータを用いることで、変異や多型の影響が少ないゲノム領域を同定し、同定したゲノム領域を避けることで gRNA 設計手法を行う。

## 4. 研究成果

当初の研究計画とほぼ同じ構想と内容の論文が、本研究課題の開始とほぼ同時に出版された (Keough et al., 2019)。そこで、ゲノム配列に着目するという構想を残したまま方向転換を行った。次に詳述する転写因子の認識配列に絞って研究課題を継続した。

転写因子認識配列の多様性を捉えることは、複雑な遺伝子発現制御機構を紐解く上で重要である。そして、公共 ChIP-seq データを統合した二次データベースの整備により、転写因子結合部位の情報を網羅的に得やすくなった。一方で、転写因子認識配列のデータベースは、位置特異的重み行列 (PWM) を用いた方法が主流であり、ゲノム上の局所的な配列情報が失われる問題があった。

本研究課題では、PWM では捉えられなかった転写因子認識配列の多様性を捉えることを目指し、ChIP-Atlas に集約された約 10 万の ChIP-seq データに対して我々が開発した MOCCS を適用した (図 1, doi:10.18908/togo2020.p013 より引用)。MOCCS では ChIP-seq データのピーク情報から、ゲノム上の塩基配列情報を損失せずに、転写因子認識配列とその結合特異性スコアを k-mer ごとに算出できる。様々な細胞型における各転写因子の認識配列とその結合特異性スコアを解析し、細胞型ごとの転写因子認識配列の多様性が明らかとなった。本研究結果で得られた知見は、転写因子を標的とした創薬など応用できる可能性があり、創薬へも貢献し得る。そして、本研究結果で得られた知見を社会発信するために、現在、生物種や細胞型による転写因子認識配列の多様性を閲覧・比較できるデータベースを構築している (図 2, 3, 4, doi:10.18908/togo2020.p013 より引用)。論文投稿の準備も行っている。本研究課題の期間終了後も、種間・細胞型間の転写因子認識配列の比較と転写調節共益因子の同定や変異の影響予測などの応用を目指す。

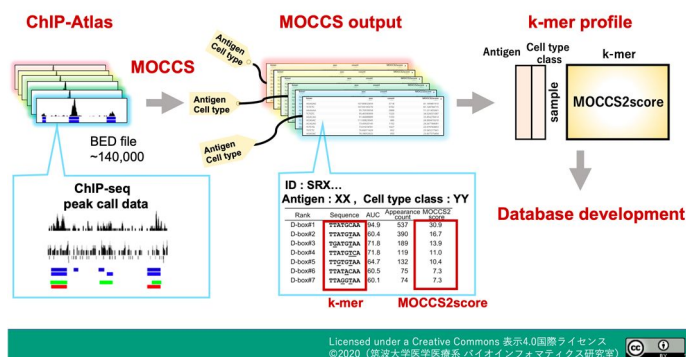
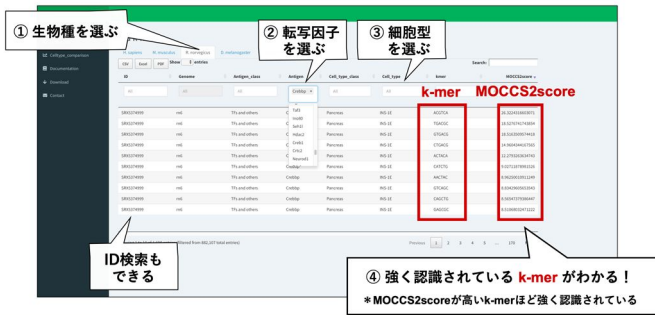


図 1: 本研究課題概要

Result

データベースの機能 - 1

興味のある転写因子の認識配列が **k-mer** としてわかる！



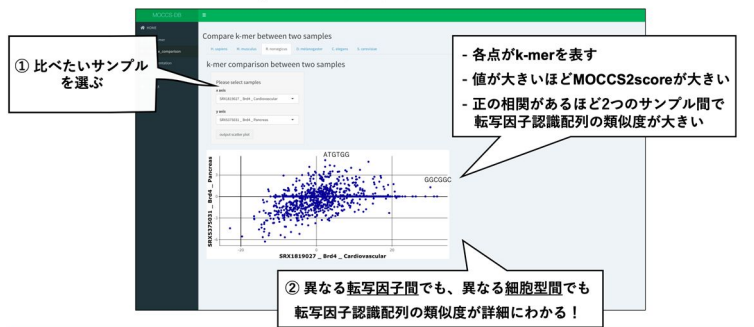
Licensed under a Creative Commons 表示4.0国際ライセンス ©2020 (筑波大学医学医療系 バイオインフォマティクス研究室)

図 2 : データベースのユーザーインターフェース例 1

Result

データベースの機能 - 2

2つのChIP-seqサンプル間で変動のある **k-mer** がわかる！

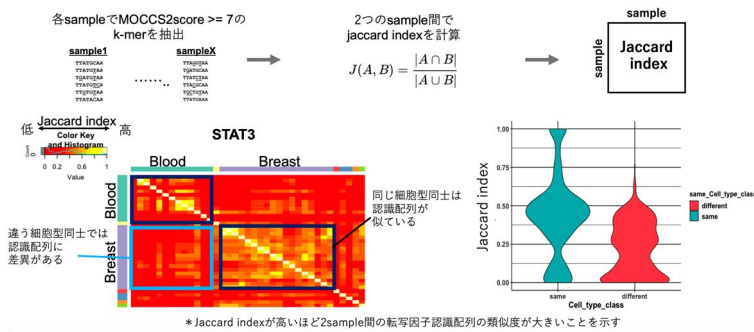


Licensed under a Creative Commons 表示4.0国際ライセンス ©2020 (筑波大学医学医療系 バイオインフォマティクス研究室)

図 3 : データベースのユーザーインターフェース例 2

Result

細胞型間で異なるk-mer profileの例  
転写因子認識配列は細胞型ごとに異なる



Licensed under a Creative Commons 表示4.0国際ライセンス ©2020 (筑波大学医学医療系 バイオインフォマティクス研究室)

図 4 : データベースのユーザーインターフェース例 3

## 5. 主な発表論文等

〔雑誌論文〕 計9件（うち査読付論文 4件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Wakida Hiroyasu, Kawata Kentaro, Yamaji Yuta, Hattori Emi, Tsuchiya Takaho, Wada Youichiro, Ozaki Haruka, Akimitsu Nobuyoshi	4. 巻 527
2. 論文標題 Stability of RNA sequences derived from the coronavirus genome in human cells	5. 発行年 2020年
3. 雑誌名 Biochemical and Biophysical Research Communications	6. 最初と最後の頁 993 ~ 999
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.bbrc.2020.05.008	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 -
1. 著者名 Kokaji T., Hatano A., Ito Y., Yugi K., Eto M., Morita K., Ohno S., Fujii M., Hironaka K., Egami R., Terakawa A., Tsuchiya T., Ozaki H., Inoue H., Uda S., Kubota H., Suzuki Y., Ikeda K., Arita M., Matsumoto M., Nakayama K.I., Hirayama A., Soga T., Kuroda S.	4. 巻 13
2. 論文標題 Transomics analysis reveals allosteric and gene regulation axes for altered hepatic glucose-responsive metabolism in obesity	5. 発行年 2020年
3. 雑誌名 Science Signaling	6. 最初と最後の頁 1236 ~ 1236
掲載論文のDOI (デジタルオブジェクト識別子) 10.1126/scisignal.aaz1236	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 -
1. 著者名 Egami R., Kokaji T., Hatano A., Yugi K., Eto M., Morita K., Ohno S., Fujii M., Hironaka K., Uematsu S., Terakawa A., Bai Y., Pan Y., Tsuchiya T., Ozaki H., Inoue H., Uda S., Kubota H., Suzuki Y., Matsumoto M., Nakayama K.I., Hirayama A., Soga T., Kuroda S.	4. 巻 24
2. 論文標題 Trans-omic analysis reveals obesity-associated dysregulation of inter-organ metabolic cycles between the liver and skeletal muscle	5. 発行年 2021年
3. 雑誌名 iScience	6. 最初と最後の頁 102217 ~ 102217
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.isci.2021.102217	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 土屋 貴穂	4. 巻 Vol.3(6)
2. 論文標題 集団ゲノムデータに基づくことによる遺伝子治療の展望	5. 発行年 2020年
3. 雑誌名 Precision Medicine	6. 最初と最後の頁 581 ~ 585
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
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 -

1. 著者名 土屋 貴穂	4. 巻 Vol.3(10)
2. 論文標題 集団ゲノムデータから展望する遺伝子治療とプレジジョンメディシンの動向	5. 発行年 2020年
3. 雑誌名 Precision Medicine	6. 最初と最後の頁 630 ~ 634
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
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 -

1. 著者名 土屋 貴穂	4. 巻 Vol.46(12)
2. 論文標題 遺伝子治療のためのゲノム編集技術と集団ゲノムデータ	5. 発行年 2020年
3. 雑誌名 Medical Science Digest	6. 最初と最後の頁 766 ~ 770
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
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 -

1. 著者名 土屋 貴穂	4. 巻 Vol.4(2)
2. 論文標題 ゲノム編集技術と大規模ゲノムデータに基づいた遺伝子治療の展望	5. 発行年 2021年
3. 雑誌名 Precision Medicine	6. 最初と最後の頁 167 ~ 171
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
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 -

1. 著者名 土屋 貴穂	4. 巻 Vol.36(5)
2. 論文標題 ゲノム編集技術と大規模ゲノムデータに基づく遺伝子治療	5. 発行年 2021年
3. 雑誌名 BIO Clinica	6. 最初と最後の頁 448 ~ 452
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
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 -

1. 著者名 内田恵理子, 平松直人, 犬飼直人, 岩井謙一, 渡辺武志, 川崎秀吉, 田村幸太郎, 土屋貴穂, 吉見英治, 高橋則彦, 伊原辰哉, 藤本和則, 山下晃人, 小野貴士, 高木観, 小野竜一, 内藤雄樹, 井上貴雄	4. 巻 Vol.50(8)
2. 論文標題 ゲノム編集技術を用いた遺伝子治療の開発動向	5. 発行年 2019年
3. 雑誌名 医薬品医療機器レギュラトリーサイエンス	6. 最初と最後の頁 pp 443-453
掲載論文のDOI (デジタルオブジェクト識別子) なし	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 -

〔学会発表〕 計5件 (うち招待講演 1件 / うち国際学会 0件)

1. 発表者名 田原 沙絵子, 土屋 貴穂, 尾崎 遼
2. 発表標題 ChIP-Atlasの二次解析による多彩な転写因子認識配列の解明および更なるデータベース化
3. 学会等名 トーゴの日 シンポジウム
4. 発表年 2020年

1. 発表者名 田原 沙絵子, 土屋 貴穂, 尾崎 遼
2. 発表標題 ChIP-seqデータベースの大規模解析で解明する細胞型ごとに多彩な転写因子認識配列
3. 学会等名 日本バイオインフォマティクス学会 2020年年会 第9回生命医薬情報学連合大会 (IBMP2020)
4. 発表年 2020年

1. 発表者名 田原 沙絵子, 土屋 貴穂, 尾崎 遼
2. 発表標題 ChIP-seqデータベースの大規模解析で解明する細胞型ごとに多彩な転写因子認識配列
3. 学会等名 生命情報科学若手の会 第12回 研究会
4. 発表年 2020年

1. 発表者名 Tsuchiya, T. and Ozaki, H.
2. 発表標題 シングルセル空間トランスクリプトームデータを用いた遺伝子制御ネットワークの空間分布解析
3. 学会等名 第8回生命医薬情報学連合大会 (IIBMP2019)
4. 発表年 2019年

1. 発表者名 土屋貴穂
2. 発表標題 System Identification of Signaling Dependent Gene Expression with Different Time-Scale Data
3. 学会等名 Institute for Frontier Life and Medical Sciences, Kyoto University, Seminar Series of Cell Fate Dynamics & Therapeutics (招待講演)
4. 発表年 2020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
---------	---------