様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科研費 KAKENHI

科学研究費助成事業　　研究成果報告書

令和　６ 年　６ 月　６ 日現在

機関番号：８２４０１
研究種目：若手研究
研究期間：2022～2023
課題番号：２２Ｋ１５３８５

研究課題名（和文）転移因子の挿入多型と疾患とのゲノムワイド関連解析

研究課題名（英文）Clarifying disease risk of mobile element variations by GWAS

研究代表者
小嶋　将平（Shohei, Kojima）

国立研究開発法人理化学研究所・生命医科学研究センター・基礎科学特別研究員

研究者番号：３０８５１７０３

交付決定額（研究期間全体）：（直接経費）　3,500,000円

研究成果の概要（和文）：本研究の成果は以下の３つである。①転移因子多型を検出するソフトウェアMEGAnEを開発しGitHubにて公開した。②転移因子多型をヒト集団スケールで解析し、世界に先駆けて転移因子多型のGWASを行った。その結果、転移因子多型が皮膚病変であるケロイドの原因バリアントの一つであり、ケロイドの発症や重症化因子であることを明らかとした。本結果をNature Genetics誌に発表した（申請者は筆頭著者かつ共責任著者）。③転移因子多型と遺伝統計解析に関するレビュー論文を執筆した。本論文はHuman Genome Variation誌に出版した（申請者は筆頭著者かつ責任著者）。

研究成果の学術的意義や社会的意義
疾患の遺伝要因の特定は、精密医療の実現の鍵となる。ゲノムワイド関連解析（GWAS）は疾患の遺伝要因を網羅的に調べる手法だが、大半のGWASは一塩基多型のみを使用しており、疾患の原因となる構造多型を見つけることができない。本研究により、ヒトの構造多型の25%を占める転移因子多型を遺伝統計解析することが可能となった。つまり、これまで見過ごされてきた疾患の遺伝要因の特定が可能となり、疾患の予防法開発と治療法開発、精密医療の推進に貢献することが期待される。実際に本研究では皮膚病変の原因変異を特定した。本結果は今後、皮膚病変リスクのより良い評価手法の開発や皮膚病変の治療開発に役立つ。

研究成果の概要（英文）：This research has made three achievements. 1) Development and open-source release of MEGAnE, a software that can precisely detect and genotype mobile element variations (MEVs). 2) GWAS of MEVs in Biobank Japan. We performed MEV GWAS and found one MEV in NEDD4 gene is one genetic cause of keloid, a skin disease. We reported this result in Nature Genetics (applicaint is the first and co-corresponding author). 3) Publishment of a review paper. I summarized statistical genetics of MEVs and published a review paper in Human Genome Variation (applicaint is the first and corresponding author).

研究分野：ゲノム医科学

キーワード：転移因子　トランスポゾン　GWAS　eQTL　遺伝統計

※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。

様　式　Ｃ－１９、Ｆ－１９－１（共通）

１．研究開始当初の背景

疾患の遺伝要因の解明は、精密医療実現に向けた重要な基礎研究である。ゲノムワイド関連解析（GWAS）は遺伝要因を網羅的に調べる手法で、バイオバンクの大規模ゲノムデータを原動力に、疾患のリスクとなるゲノム領域や変異が多く同定されている。しかし、これまでは一塩基多型や短鎖の配列挿入・欠損のみが GWAS に使用されており、長鎖の配列挿入・欠損である構造多型は用いられてこなかった。つまり、疾患の原因となる構造多型の同定はあまり行われておらず、見過ごされてきた疾患リスクである。

転移因子（トランスポゾン）は、ゲノム内で自身の移動する DNA 配列である。現在でも転移活性のあるヒト転移因子のファミリーとして LINE-1、SVA、*Alu* が知られており、これらは通称「コピー・アンド・ペースト」と呼ばれる方法で転移する。進化的に新しい転移は、ヒトゲノム間で多型となる。LINE-1、SVA、*Alu* の配列長は、それぞれおよそ 6kb, 2kb, 0.3kb であり、転移因子多型は、構造多型の一角を担う。実際、ヒトゲノムに見つかる構造多型のおよそ 25%は転移因子多型である。

転移因子多型は全ゲノムシーケンス（WGS）データから検出可能である。例えば、1000 人ゲノムプロジェクトが開発した MELT と呼ばれるソフトウェアが使用可能である。しかし、検出とジェノタイピングの精度が良くなく、バイオバンクなどのヒト集団スケールでの解析には用いられていない。つまり、バイオバンクにおいて転移因子多型を調べることはこれまで難しく、GWAS や eQTL 解析などの遺伝統計解析を用いた疾患リスクの評価が行われていなかった。

２．研究の目的

転移因子多型を高精度に検出するソフトウェアの開発しバイオバンクの疾患ゲノムへの適用、ヒト集団スケールで転移因子多型を同定するとともに、遺伝統計解析に転移因子多型を取り入れ GWAS により疾患リスクを評価することを目的とした。

３．研究の方法

◆ 転移因子多型を同定するソフトウェアの開発

短鎖のイルミナ全ゲノムシーケンスデータを読み込み、転移因子の挿入を同定するアルゴリズムを開発した。数万人規模の解析を想定し、実行速度の速い C++でコアアルゴリズムを実装した。１ゲノムあたり 1 時間ほどで実行可能なソフトウェアを目指した。ソフトウェアの精度の検証として、35 人の長鎖全ゲノムシーケンスデータを用いるとともに、100 箇所の挿入を 2000 人において PCR で増幅し、ジェノタイピングの精度を検証した。

◆ヒト集団スケールでの転移因子多型の同定

開発したソフトウェアをバイオバンクに適用し、ヒト集団スケールで転移因子多型を同定した。具体的には、以下のバイオバンクに適用した。特に、1000 人ゲノムプロジェクトのデータに適用、多様なヒトにおける転移因子の挿入をカタログし、ジェノタイプ・インピューテーションの参照パネルを作成した。

データベース	解析したゲノム数	ヒト集団
1000 人ゲノムプロジェクト	3,222	グローバル
バイオバンク・ジャパン	4,880	日本人集団
SFARI	9,201	主に欧米人集団
GTEx	838	主に欧米人集団

◆GWAS による疾患リスクの評価

バイオバンク・ジャパンでは 18 万人のゲノムを SNP アレイで解析している。本研究では、この 18 万人における転移因子多型をインピューテーションし、GWAS に用いた。GWAS には一塩基多型と転移因子多型の両方を用い、42 疾患との関連を解析した。

４．研究成果

◆ 転移因子が駆動するゲノム進化：東アジア人集団では転移因子多型が多い

1000 人ゲノムプロジェクトのデータの解析から、東アジア人集団に LINE-1 と SVA の転移因子多型が多いことが明らかとなった（図 1）。これは、東アジア人集団のゲノムの進化の原動力として、他のヒト集団よりも転移因子が大きく寄与していることを示唆する。アフリカ人集団においても多いが、これはアフリカ人集団の遺伝的多様性を反映しているからである。東アジア人集団において転移因子多型が多い理由は不明であるが、転移活性のある転移因子のコピー数が多いことや、転移を抑制する因子の活性が低いことが考えられる。

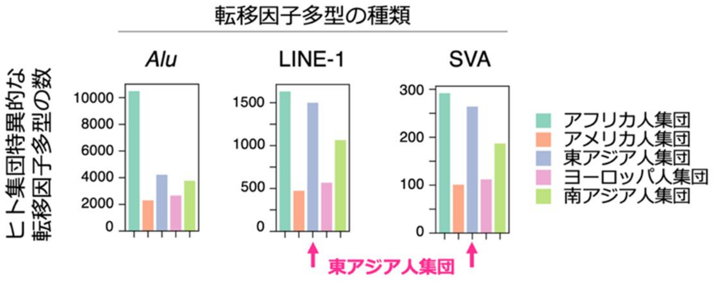

図１. ヒト集団に特異的に見つかった転移因子多型の数

◆ 転移因子多型の疾患リスク：ケロイドの原因バリアント

バイオバンク・ジャパンを活用した GWAS により、皮膚病変であるケロイドの原因バリアントとして転移因子多型を同定した（図 2）。ケロイドの GWAS を行った結果、5 つのリスク領域を同定した。これは過去の GWAS と同じ結果である。しかし、一つのリスク領域に LINE-1 の挿入多型が含まれていた。LINE-1 挿入は、ケロイドの原因遺伝子の一つである *NEDD4* のイントロンに存在していた。長鎖シーケンスにより、このリスクハプロタイプには LINE-1 挿入以外の構造多型がないこと、および *NEDD4* のミスセンス変異がないことから、LINE-1 の挿入が原因バリアントである可能性が考えられた。特に、LINE-1 には RNA ポリメラーゼの転写開始点が存在することから、*NEDD4* のエンハンサーとして機能することが考えられた。

LINE-1 挿入が原因変異であるか、および LINE-1 が *NEDD4* のエンハンサーとして機能するかを検証するため、iPS 細胞由来の皮膚細胞（線維芽細胞）にて LINE-1 挿入をノックアウトし、*NEDD4* への影響を評価した。その結果、LINE-1 は *NEDD4* のエンハンサーであることがわかった。つまり、LINE-1 挿入は *NEDD4* の転写を増加させ、ケロイドのリスクとなることが明らかとなった。

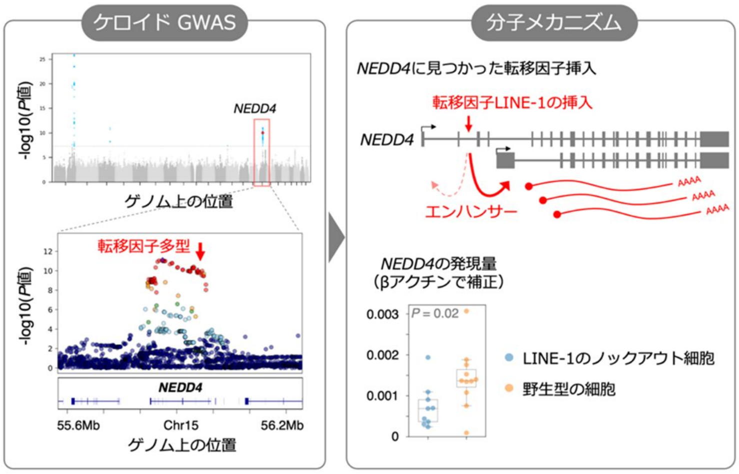

図２. LINE-1 多型は *NEDD4* の発現を増加させることで、ケロイド発症・重症化のリスクとなる

◆ 論文

これらの研究成果を *Nature Genetics* 誌に報告し（研究代表者が筆頭著者かつ共責任著者）、また *Human Genome Variation* 誌にレビュー論文を発表した（研究代表者が筆頭著者かつ責任著者）。

５．主な発表論文等

〔雑誌論文〕　計4件（うち査読付論文　3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	4．巻
Kojima Shohei, Koyama Satoshi, Ka Mirei, Saito Yuka, Parrish Erica H., Endo Mikiko, Takata Sadaaki, Mizukoshi Misaki, Hikino Keiko, Takeda Atsushi, Gelinas Asami F., Heaton Steven M., Koide Rie, Kamada Anselmo J., Noguchi Michiya, et al.	55
2．論文標題	5．発行年
Mobile element variation contributes to population-specific genome diversification, gene regulation and disease risk	2023年
3．雑誌名	6．最初と最後の頁
Nature Genetics	939～951
掲載論文のDOI（デジタルオブジェクト識別子）	査読の有無
10.1038/s41588-023-01390-2	有
オープンアクセス	国際共著
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	－

1．著者名	4．巻
Tamura Kenji、Kanazashi Yuki、Kawada Chiaki、Sekine Yuya、Maejima Kazuhiro、Ashida Shingo、Karashima Takashi、Kojima Shohei、Parrish Nickolas F、Kosugi Shunichi、Terao Chikashi、Sasagawa Shota、Fujita Masashi、Johnson Todd A、Momozawa Yukihide、Inoue Keiji、Shuin Taro、Nakagawa Hidewaki	32
2．論文標題	5．発行年
Variant spectrum of von Hippel-Lindau (VHL) disease and its genomic heterogeneity in Japan	2023年
3．雑誌名	6．最初と最後の頁
Human Molecular Genetics	2046～2054
掲載論文のDOI（デジタルオブジェクト識別子）	査読の有無
10.1093/hmg/ddad039	有
オープンアクセス	国際共著
オープンアクセスとしている（また、その予定である）	－

1．著者名	4．巻
Kojima Shohei	-
2．論文標題	5．発行年
Preparation of Non-overlapping Transposable Elements (TEs) Annotation by Interval Tree	2022年
3．雑誌名	6．最初と最後の頁
Methods in Molecular Biology	353～360
掲載論文のDOI（デジタルオブジェクト識別子）	査読の有無
10.1007/978-1-0716-2380-0_21	無
オープンアクセス	国際共著
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	－

1．著者名	4．巻
Kojima Shohei	11
2．論文標題	5．発行年
Investigating mobile element variations by statistical genetics	2024年
3．雑誌名	6．最初と最後の頁
Human Genome Variation	-
掲載論文のDOI（デジタルオブジェクト識別子）	査読の有無
10.1038/s41439-024-00280-1	有
オープンアクセス	国際共著
オープンアクセスとしている（また、その予定である）	－

〔学会発表〕　計4件（うち招待講演　3件／うち国際学会　2件）

１．発表者名
Shohei Kojima, Nicholas F. Parrish

２．発表標題
Mobile element variation drives population-specific genome diversification, gene regulation, and disease risk

３．学会等名
Cold Spring Harbor Laboratory Meeting: Transposable Elements（国際学会）

４．発表年
2022年

１．発表者名
Shohei Kojima

２．発表標題
Polymorphisms of endogenous retroviruses in humans: from detection methods to disease risk

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
小嶋 将平、Nicholas F. Parrish

２．発表標題
ヒトゲノムの構成成分からゲノム機能と疾患を読み解く

３．学会等名
日本遺伝学会第94回大会（招待講演）

４．発表年
2022年

１．発表者名
小嶋 将平、Nicholas F. Parrish

２．発表標題
Mobile element variation contributes to population-specific genome diversification, gene regulation, and disease risk

３．学会等名
Human Genetics Asia 2023（招待講演）（国際学会）

４．発表年
2023年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

転移因子多型の疾患リスクを解明
https://www.riken.jp/press/2023/20230512_1/index.html

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