

令和 6 年 5 月 22 日現在

機関番号：17102

研究種目：基盤研究(B)（一般）

研究期間：2021～2023

課題番号：21H02681

研究課題名（和文）日本人集団ゲノムワイドクロマチン相互作用情報と構造多型との関連解析と疾患要因探索

研究課題名（英文）Genome-Wide Analysis of Chromatin Interactions and Structural Variants in the Japanese Population Towards the Discovery of Disease-Related Factors

研究代表者

長崎 正朗（Nagasaki, Masao）

九州大学・生体防御医学研究所・教授

研究者番号：90396862

交付決定額（研究期間全体）：（直接経費） 13,700,000円

研究成果の概要（和文）：ゲノムワイドクロマチンアクセシビリティや相互作用プロファイルが取得可能な短鎖型測定技術と、長鎖型シーケンサによって得られる構造多型プロファイルとを組み合わせることで、日本人集団におけるゲノムワイドクロマチンアクセシビリティの差異の検討を行いながら、同集団中に含まれる一塩基多型および構造多型のうちどの多型がこの染色体上の各クロマチン相互作用にどのように影響を与えるのか研究を推進した。この成果の1例としてLILR領域の新規融合遺伝子の同定と機能解析を行うことができた。また発展的な取り組みとしてJoGoデータベース開発を開始するきっかけとなった (<https://jogo.csml.org/>)。

研究成果の学術的意義や社会的意義

この研究では、クロマチンアクセシビリティと長鎖型シーケンシングを組み合わせ、日本人集団の染色体上の構造多型の特定やクロマチン相互作用の違いを解析しました。研究期間内に新規のLILR領域の融合遺伝子を特定し、その影響を解明することができました。同研究基盤をさらに推進することで、他の遺伝子の遺伝学的多様性の理解を深めるとともに、将来の医学研究において重要なリソースの1つとなります。

研究成果の概要（英文）：By combining measurement technologies that capture genome-wide chromatin accessibility and interaction profiles from short-read sequencing and whole-genome long-read sequencing, which provides structural polymorphism profiles, this research advanced the study of the differences in genome-wide chromatin accessibility among the Japanese population. It also explored how specific polymorphisms -including single nucleotide polymorphisms and structural variants- affect chromatin interactions on chromosomes. One outcome of this research was identifying and functionalizing a novel fusion gene in the Leukocyte Immunoglobulin-Like Receptor (LILR) region (Hirayasu et al. *Frontiers in Immunology* 2023). Furthermore, this work has led to the development of the key database development concept of the JoGo database in future (<https://jogo.csml.org/>).

研究分野：バイオインフォマティクス

キーワード：構造多型 長鎖型シーケンサ 大規模ゲノム情報解析 クロマチン構造

様式 C-19、F-19-1 (共通)

1. 研究開始当初の背景

DNAの核内配置が転写制御、複製、DNA修復など影響を与えることが知られている。一方、長鎖型シークエンサの開発が進み、短鎖型シークエンサでは困難であった全ゲノムレベルでの構造多型のプロファイルの取得が可能となっている。そこでゲノムワイドクロマチン相互作用計測技術情報と長鎖型シークエンサによって得られる構造多型プロファイルを組み合わせ、日本人集団におけるゲノムワイドクロマチン相互作用の差異の検討をすると共に、同集団中に含まれる頻度5%以上の一塩基多型、および、構造多型のうちどの多型がこの染色体上の各クロマチン相互作用にどのように影響を与えるのか、また、その多型が疾患にどう影響をするか研究を行う。

2. 研究の目的

DNAの核内の配置が転写制御、複製、DNA修復など色々な影響を与えること、また、構造は細胞種特異的であることが知られている。一方、長鎖型シークエンサの開発が進み、短鎖型シークエンサでは困難であった全ゲノムレベルでの構造多型のプロファイルの取得が可能となっている。また、クロマチン相互作用は遺伝的な要因によって個人差があることが知られている。さらに、その差が疾患に影響を与える報告もある。そこでゲノムワイドクロマチン相互作用を取得することができる技術と長鎖型シークエンサによって得られる構造多型プロファイルを組み合わせ、日本人集団におけるゲノムワイドクロマチン相互作用の差異の検討をするとともに、同集団中に含まれる頻度5%以上の一塩基多型、および、構造多型のうちどの多型がこの染色体上の各クロマチン相互作用にどのように影響を与えるのか、また、その多型が疾患にどのように影響をするかについて研究を行う。

3. 研究の方法

【令和3年度】

【課題1】大川が試験的に検証を行った実験プロトコルに基づいて、凍結保存をしている不死化B細胞株を適宜細胞培養し、50検体を目標にゲノムワイドクロマチン相互作用ライブラリの構築を行う(100アレルで日本人集団中で頻度5%以上の変異の約90%を3回以上計測される数)。さらに、大川が管理するNovaSeqを用いて大規模シークエンス情報を取得する。得られた情報について、長崎が情報解析ツールを用いて、ゲノムワイドクロマチン構造プロファイルの取得を行う。さらに、日本人集団内でのプロファイル比較を行う。

【課題2】河合と長崎が試験的に構築を行った長鎖型シークエンサの構造多型プロファイルの同定のための解析パイプラインを実行することで、【課題1】のデータと統合解析をするための構造多型プロファイルを取得する。

【課題3】国際1000人プロジェクトなどで公開されているSNVなどの情報を用いて、【課題1】で得られた構造多型プロファイルとの関連解析を着手する。

【課題4】ヒトの希少疾患を中心とした変異、構造多型との関係を整理した国際データベースClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>)、および、ヒト疾患・形質と変異との関連解析を収集した国際データベースGWASCatalog (<https://www.ebi.ac.uk/gwas/>)に登録されている変異情報を整備することで、疾患(身長、体重などの形質も含む)データベースとの比較ができる体制を整えることに着手する。

【令和4年度】

【課題1, 2】前年度に続き解析パイプラインの精度向上を他の課題の解析結果をフィードバックすることで必要に応じて行う。

【課題3】令和3年度の関連解析の結果について整理をすすめ【課題6】のデータベースとして利用できる体制を進める。

【課題4】ホモ欠失を中心とした構造多型を重点的に検証する(ヘテロ欠失に比べてクロマチン相互作用の変化の影響が把握しやすいため)。その後、ヘテロ欠失に対しての影響がホモ欠失で想定されるプロファイルの変化を含んだ変化になっているかを検討することで、ヘテロ欠失か集団中で観測されていない欠失についても拡大し、クロマチン相互作用の評価ができるようにモデルの拡張を進める。同成果を【課題6】のデータベースとして利用できる体制を進める。

【課題5】【課題2】で得られたプロファイルのうち、逆位およびコピー数多型とクロマチン相互作用の変化についての調査に着手する。

【課題6】【課題3】【課題4】の結果をデータベースとしてまとめる。

【課題7】【課題6】のデータベースと昨年度に進めていたGWASで得られる変異情報、希少疾患で得られる変異情報との関係を比較することで、クロマチン相互作用に影響を与えるSNV、構造多型との比較調査を行う。また、興味深いクロマチン相互作用の変化については、必要に応じて

て、大川と実験的にトランスクリプトームや他のレイヤーでの細胞の状態を取得することで追試するなどの検討を開始する。

【令和5年度】

【課題1-2】前年度に続き解析パイプラインの精度向上を他の課題の解析結果をフィードバックすることで必要に応じて行う。

【課題3-4】前年度の解析結果をもとに必要に応じて解析手法を改善し精度の向上を行う。

【課題5】【課題3】【課題4】【課題5】の情報整理を進める。

【課題6】前年度に加え【課題5】で得られる複雑な構造多型と希少疾患、多因子疾患、ヒト形質との関係性について調査を行う。

4. 研究成果

ゲノムワイドクロマチンアクセシビリティや相互作用プロファイルを取得することができる短鎖型測定技術と、長鎖型シーケンサによって得られる構造多型プロファイルとを組み合わせることで、日本人集団におけるゲノムワイドクロマチンアクセシビリティの差異の検討をすすめた。特に、当初予定ではHiCで50検体を目標としていたが人数規模を確保するために100検体のATAC-Seqによるクロマチンのアクセシビリティの情報取得を分担研究者の大川と進めた。

このATAC-Seqの情報を用いることで、ATAC-QTL解析を行うことが可能となり、どのSNPや構造多型がどのATAC-Seqのピークに影響を与えるか探索を行うことができる情報整理を行うことができた。これにより、同集団中に含まれる一塩基多型、および、構造多型のうちどの多型がこの染色体上の各クロマチン相互作用にどのように影響を与えるのかGWASで得られる変異情報と統合して解析できるようになった。

この情報基盤を活用した成果の1例として、LILR領域の新規融合遺伝子の同定と機能解析を行う情報基盤の1つとして得ることができた (Frontiers in Immunology, Hirayasu *et al.* 2024)。他に新たな一部のヒト集団が有する新規遺伝子候補を同基盤で同定するなどしている (in preparation)。

この研究で得られた情報について、さらに発展的な取り組みとして、JoGoデータベース開発を開始するための情報基盤の1つとして活用することができた (<https://jogo.csml.org/>)。

今後、同情報基盤をさらに拡充発展させることで、他の遺伝子の遺伝学的多様性の理解を深めるとともに、将来の医学研究において重要なリソースの1つとして活用できると考えている。

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 K. Hirayasu, S.S. Khor, Y. Kawai, M. Shimada, Y. Omae, G. Hasegawa, Y. Hashikawa, H. Tanimoto, J. Ohashi, K. Hosomichi, A. Tajima, H. Nakamura, M. Nakamura, K. Tokunaga, R. Hanayama, M. Nagasaki	4. 巻 15
2. 論文標題 Identification of the hybrid gene LILRB5-3 by long-read sequencing and implication of its novel signaling function	5. 発行年 2024年
3. 雑誌名 Frontiers in Immunology	6. 最初と最後の頁 1398935
掲載論文のDOI（デジタルオブジェクト識別子） 10.3389/fimmu.2024.1398935	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 Y. Mori, K. Ueno, D. Chiba, K. Hashimoto, Y. Kawai, K. Baba, H. Tanaka, T. Aki, M. Ogasawara, N. Shibasaki, K. Tokunaga, T. Aizawa, M. Nagasaki	4. 巻 24
2. 論文標題 Genome-Wide Association Study and Transcriptome of Japanese Patients with Developmental Dysplasia of the Hip Demonstrates an Association with the Ferroptosis Signaling Pathway	5. 発行年 2023年
3. 雑誌名 International Journal of Molecular Sciences	6. 最初と最後の頁 5019
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/ijms24055019	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7件（うち招待講演 3件/うち国際学会 0件）

1. 発表者名 長崎 正朗
2. 発表標題 未来医療へのヒト情報解析基盤構築と実装
3. 学会等名 未来社会デザイン統括本部 & データ駆動イノベーション推進本部 合同シンポジウム2023
4. 発表年 2023年

1. 発表者名 長崎 正朗
2. 発表標題 ヒトゲノム情報や臨床情報のセキュリティと情報解析の取り組みについて
3. 学会等名 xSC2023Q セキュリティとスーパーコンピュータシンポジウム（招待講演）
4. 発表年 2023年

1. 発表者名 長崎 正朗
2. 発表標題 ヒトゲノムと臨床情報の統合解析に向けたハイブリッドクラウド基盤構築とパブリッククラウドの活用
3. 学会等名 教育と研究のためのDXフォーラム（招待講演）
4. 発表年 2023年

1. 発表者名 長崎 正朗
2. 発表標題 ハイブリッドクラウドを用いたゲノム情報に基づく構造多型パネルの構築とアノテーション
3. 学会等名 学際大規模情報基盤共同利用・共同研究拠点 第15回シンポジウム
4. 発表年 2023年

1. 発表者名 長崎 正朗
2. 発表標題 ハイブリッドクラウド構築とゲノム情報解析の効率的な運用に関する研究
3. 学会等名 学際大規模情報基盤共同利用・共同研究拠点 第14回シンポジウム
4. 発表年 2022年

1. 発表者名 長崎 正朗
2. 発表標題 日本人の公開可能な長鎖型集団パネル構築とその意義について
3. 学会等名 PacBio ユーザーグループミーティング 2022（招待講演）
4. 発表年 2022年

1. 発表者名 長崎 正朗
2. 発表標題 ハイブリッドクラウド構築とゲノム情報解析の効率的な運用に関する研究
3. 学会等名 学際大規模情報基盤共同利用・共同研究拠点 第13回シンポジウム
4. 発表年 2021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
研究 分担者	河合 洋介 (Kawai Yosuke) (30435515)	国立研究開発法人国立国際医療研究センター・その他部局 等・副プロジェクト長 (82610)	
研究 分担者	大川 恭行 (Ohkawa Yasuyuki) (80448430)	九州大学・生体防御医学研究所・教授 (17102)	

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		
米国	The Consortium of Long Read Sequencing		