

令和 2 年 6 月 15 日現在

機関番号：34417

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K07255

研究課題名(和文)疾患関連変異の探索に適した高次元型日本人ゲノム参照配列の構築

研究課題名(英文) Construction of high-dimensional Japanese genome reference sequences suitable for the search for disease-associated mutations

研究代表者

日笠 幸一郎 (HIGASA, Koichiro)

関西医科大学・医学部・教授

研究者番号：10419583

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：現行のヒト参照ゲノムは、直線的なハプロイドDNA配列である。ヒトゲノムには参照ゲノムと比較して何百万もの変異が存在するため、この参照構造は、ヒト集団における遺伝的多様性の解析に実用的な限界をもたらしている。本課題では、日本人3,135人の全ゲノム配列データから抽出した複雑な遺伝的構成要素を、グラフ構造として現行のヒト参照配列に統合し、多様な変異をより正確に同定可能な汎用型ヒトゲノム参照配列を構築に成功した。蓄積される遺伝的変異をヒトゲノム参照配列に統合することは、精密医療に向けた病原性変異の解釈に適した最先端の参照構造を開発するための有望な戦略である。

研究成果の学術的意義や社会的意義

現在、精密医療を目指し世界中で進められている大規模ゲノム解析は、欧米人を対象とした情報の収集に偏っている。本課題では、日本人の遺伝的多様性を網羅する3,135人を対象に全ゲノム配列解析を行い、44,757,785箇所の変異を検出した。そのうちアレル頻度1%未満の98.9%の変異が今まで報告のない変異であることが判明した。またこれらの情報を活用することにより、正確な遺伝子型の推定や変異の同定が可能であることが分かった。本知見は、民族的に多様な集団からの遺伝的変異を既存のゲノム情報に統合することが、あらゆる人種を対象とした精密医療を実現するために不可欠であることを示唆するものである。

研究成果の概要(英文)：The current human reference genome is a linear haploid DNA sequence. This reference structure poses practical limitations due to the prevalence of genetic diversity in human populations because any given human genome has millions of variations compared with the reference genome. Therefore, a strategy to reconstruct a more complete version of the reference sequence will be essential. To construct a state-of-the-art reference sequence, we elucidated the complex genetic components from whole genome-sequencing data of 3,135 Japanese individuals and integrated them into the current reference sequence as a graph structure. The Japanese graph reference sequence significantly contributed to the precise characterization of variants. The integration of accumulating knowledge of genetic variation into the current reference sequence is a promising strategy to develop a leading-edge reference structure appropriate for interpretation of pathogenic variants towards precision medicine.

研究分野：ゲノム解析

キーワード：ヒトゲノム構造多様性

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

国際ヒトゲノム計画によって“ヒトゲノム参照配列”が解読されて以来、多種多様なゲノム上の変異(バリエーション)が報告され、想像以上に我々個体間のゲノム配列は異なる事が明らかとなってきた。その主たる要因は、国際 HapMap 計画で解析された一塩基変異だけでなく、構造変異と称されるより大きな変異や未知のゲノム領域の存在である。こうした知見は、関連解析・連鎖解析等を用いて疾患原因変異を同定する際に、個々の解析症例に最適な参照配列情報が、比較対照として必要不可欠である事を明確に意味する。現行のヒトゲノム参照配列は、少数の西欧人由来であることから、様々な人種を対象とした疾患関連変異探索研究において、必要とされる精度の変異情報が得られない状況にある。ヒトゲノムの可変性は一次元の情報として格納するには極めて複雑かつ限界であり、ヒトゲノム参照配列の有り方を根本的に改める時期に来ている。こうした現状の中、先行研究として我々日本人のゲノム解析に最適な参照配列を新規に構築する試みは意義深い。最近まで人種毎の参照配列の新規構築が困難であった理由は、ゲノム上に散在する変異の多様性と特徴的配列類似性に、既存の解析機器と方法論が対応できていなかった点が挙げられる。

2. 研究の目的

本研究課題では一分子バーコード技術と希少なヒトハプロイドゲノム検体を用いることにより高精度の日本人ゲノム参照配列を新規構築する。更に、構築した参照配列を鋳型として、網羅的ゲノム変異情報を統合した高次元型の参照配列に拡張し、新たに提案する“個別化ゲノム解析法”を用いた効率的な病因変異の探索に活用することを目的とする。本課題の達成により、ゲノム研究者はもとより、様々な医療関係者にとって極めて有用な情報基盤と方法論が確立される。本研究の戦略は他の人種にも応用可能であり、将来のヒトゲノム解析の方向性を示す重要な先導研究となり得る。

3. 研究の方法

本課題では、日本人由来の「全胞状奇胎(ハプロイドゲノム検体)」と「大規模全ゲノム解析情報」を用い、ゲノム医学研究に必須の高精度・高次元日本人ゲノム参照配列を完成させる。2017年度は、1分子バーコード技術を用いて、全胞状奇胎から高精度の一次元日本人ゲノム参照配列を構築し、以降は2017年度に構築した全胞状奇胎由来の一次元参照配列を鋳型として、大規模全ゲノム解析より得られた多様な変異情報を統合し、参照配列を高次元(グラフ構造)化する。原因未同定の遺伝性疾患について、構築した高次元型の参照配列情報を活用した変異同定“個別化ゲノム解析”を試み、その有用性を評価する。

4. 研究成果

本課題における主な成果として、ゲノム解析によって同定した良性成人型家族性ミオクローヌスてんかんの新たな原因遺伝子についてその概要を報告する。

(1) 背景

良性成人型家族性ミオクローヌスてんかん(benign adult familial myoclonic epilepsy, BAFME)は、日本において、1990年頃に疾患概念が確立された疾患で、厚生労働省が定める指定難病の一つである進行性ミオクローヌスてんかんに含まれる疾患である。症状は、手指のふるえとてんかんであり、発症年齢の多くは20~60歳で、遺伝形式は常染色体優性遺伝形式である。本疾患の報告は、大部分が国内からなされており、日本に多く見られると考えられている。てんかんの頻度はそれほど高くはなく、抗てんかん薬によってコントロールされる。細かい手のふるえ(振戦様ミオクローヌスとも呼ばれる)については、ほとんど変わらないが、非常に緩徐にはあるが徐々に増悪することがある。通常、認知機能障害などは出現しないこと、進行はあっても極めて緩徐であることから、良性という名前が付けられている。この疾患については、日本で多くの研究者により、精力的な研究が進められてきており、その遺伝子座が第8番染色体長腕にあることが判明していたが、候補領域の全遺伝子のエクソンの詳細な解析を行っても、原因となる遺伝子変異は見出せておらず、その原因は不明であった。

(2) 今回の成果

今回、多施設共同研究により、51家系の方々よりご協力を頂き、SAMD12遺伝子のイントロンにTTTCA・TTTTA繰り返し配列の異常伸長変異を発見した。連鎖解析という手法で、本疾患の遺伝子座が第8番染色体長腕に存在することを確認した後、さらにこの領域の多型マーカーについて詳しく調べた。その結果、非常に狭い領域で、発症者が共通して、特定の多型パターンの組み合わせを持っていることを発見し、候補領域を約420万塩基対から13万塩基対の範囲に大幅に狭めることに成功した。次世代シーケンサーを用いて、ゲノム全域の配列を決定したが、候補領域内の遺伝子のエクソン領域(翻訳領域)には、疾患発症の原因となる変異を見出すことができなかった。次に、イントロン領域に存在する変異を検討したところ、SAMD12遺伝子のイントロン4に、本来TTTTAという5塩基が20回程度短く繰り返す配列が存在する部位において、TTTTA繰り返し配列の異常伸長が認められたのみならず、通常は存在しないTTTCAという5塩基の繰り

返し配列が異常伸長している配列が挿入されていることを発見した。また、1万塩基対以上を解読できる最先端の1分子シーケンサーを用いた解析で、挿入のされ方には2通り見られることがわかった。この繰り返し配列の異常伸長変異は、51家系中49家系の発症者の方全てで認められた(図)。

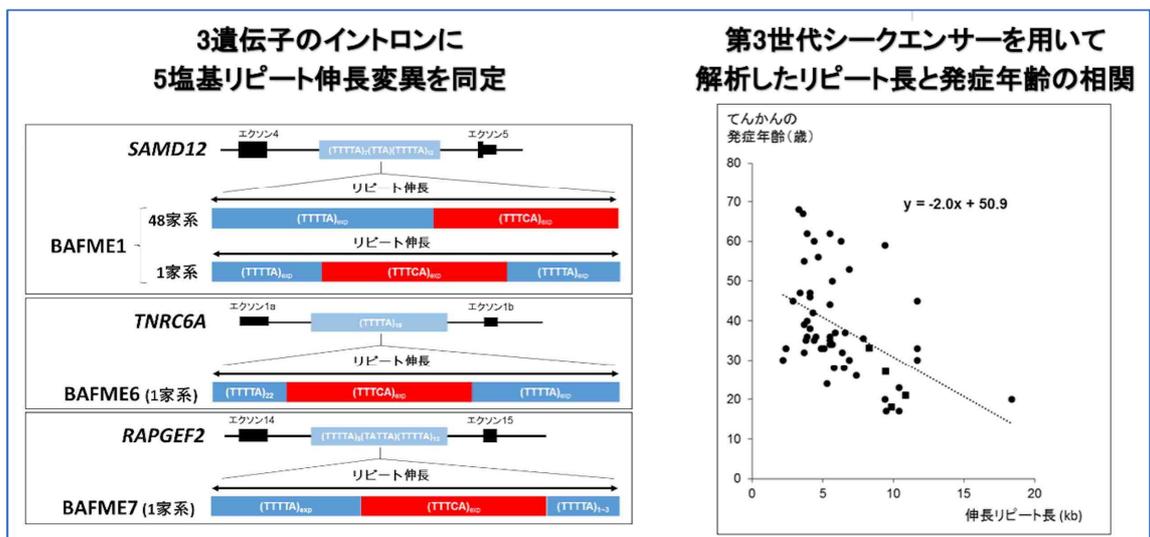
繰り返し配列の長さは、患者において合計440回~3680回(2,200~18,400塩基)に及び、この長さはてんかんやミオクローヌスの発症年齢と有意に逆相関を認めた(図)。また、世代を経るとともに、繰り返し配列の長さは長くなる傾向があり、これまでの臨床的観察により認められている表現促進現象と対応することを見いだした。健常者においては稀に、TTTTAの繰り返し配列の伸長配列が見いだされたことから、TTTCA繰り返し配列がイントロンに挿入されることが、本疾患の病態の形成に主要な役割を果たしていると考えた。

別の2つの遺伝子のイントロンにも同様のTTTCA・TTTTA繰り返し配列の異常伸長変異を発見した。SAMD12遺伝子のリピート伸長変異は、患者の96%で認められる、最も多い変異であることがわかったが、残る2家系においては、SAMD12遺伝子に異常を認められなかった。全ゲノム配列解析のデータからリピート伸長変異を見いだすために開発されたTRhistプログラムを用いて、繰り返し配列の異常伸長変異の探索を試みた。その結果、1家系ではTNRC6A遺伝子に、もう1家系ではRAPGEF2遺伝子のイントロン内に、SAMD12遺伝子に見られたものと同様のTTTTA繰り返し配列が異常伸長しているのみならず、内部にTTTCA繰り返し配列が異常伸長して挿入されていることを発見した。この2疾患をそれぞれ、BAFME6型(BAFME6)、7型(BAFME7)とした(図)。

今回の研究で、3つの遺伝子のイントロンに、同一のTTTCA繰り返し配列の異常伸長配列が挿入されることで、進行性ミオクローヌステんかんという疾患がもたらされることが明らかとなった。挿入部位はイントロンで、タンパク質を作る情報を持った(タンパク質をコードする)領域ではないこと、挿入されている遺伝子は異なっても同一の繰り返し配列の異常伸長が同一の疾患を引き起こすことから、繰り返し配列そのものが重要で、そこから転写されて生成されるRNA分子が、神経細胞に対して機能障害をもたらすものと考えられた。これを裏付けるように、SAMD12遺伝子に繰り返し配列の異常伸長を持つ患者では、神経細胞内にTTTCA繰り返し配列がRNAに転写され、神経細胞の核の中でRNA fociという異常凝集体を形成していることを見いだした。このことは、RNAによる病態機序を強く支持するものと考えられる。

(3) 本研究の意義と今後の展望

これまで発症の機構が不明であった本疾患の発症機構が解明できたことから、遺伝子検査により本疾患の診断確定ができるようになり、これまでに経験的に蓄積されてきている、本疾患の臨床的特徴や治療法等に関する正確な情報を患者に提供できるようになる。これまでの治療法は、一般的に用いられている抗てんかん薬であるが(対症療法)、原因が究明されたことから、より効果的な治療法の開発研究が発展すると期待される(病態機序を直接治療する治療薬の開発)。てんかんで見出された新しい発症機構であり、他のてんかんの原因究明にも大きく貢献すると期待される。また、非コード領域の繰り返し配列の異常伸長変異を探索することは難しいため研究が進んでいない部分もあるが、本研究により、従来考えられてこなかったような多くの疾患についても、繰り返し配列がその原因となっている可能性を示唆するものとする。複数の遺伝子で、同じ繰り返し配列の異常伸長を見だし、異常の存在する遺伝子によらず、繰り返し配列の異常伸長そのものが疾患発症に直接関与することを見出した初めての報告でもあり、次世代シーケンサーを駆使した研究手法を含めて、今後、さらに数多くの神経疾患で、その原因究明に大きく貢献すると期待される。



5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 6件/うち国際共著 1件/うちオープンアクセス 4件）

1. 著者名 Anderson-Trocme Luke, Farouni Rick, Bourgey Mathieu, Kamatani Yoichiro, Higasa Koichiro, Seo Jeong-Sun, Kim Changhoon, Matsuda Fumihiko, Gravel Simon	4. 巻 37
2. 論文標題 Legacy Data Confound Genomics Studies	5. 発行年 2019年
3. 雑誌名 Molecular Biology and Evolution	6. 最初と最後の頁 2~10
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/molbev/msz201	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 該当する
1. 著者名 Furusawa Yoshihiko, Yamaguchi Izumi, Yagishita Naoko, Tanzawa Kazumasa, Matsuda Fumihiko, Yamano Yoshihisa, RADDAR-J Research and Development Group	4. 巻 3
2. 論文標題 National platform for Rare Diseases Data Registry of Japan	5. 発行年 2019年
3. 雑誌名 Learning Health Systems	6. 最初と最後の頁 e10080 ~ e10080
掲載論文のDOI（デジタルオブジェクト識別子） 10.1002/lrh2.10080	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 Ishiura Hiroyuki, Doi Koichiro, Mitsui Jun, Yoshimura Jun, Matsukawa Miho Kawabe, Fujiyama Asao, Toyoshima Yasuko, Kakita Akiyoshi, Takahashi Hitoshi, Suzuki Yutaka, Sugano Sumio, Qu Wei, Ichikawa Kazuki, Yurino Hideaki, Higasa Koichiro, Shibata Shota, Mitsue Aki, Tanaka Masaki, Ichikawa Yaeko, Takahashi Yuji et al.	4. 巻 50
2. 論文標題 Expansions of intronic TTCA and TTTA repeats in benign adult familial myoclonic epilepsy	5. 発行年 2018年
3. 雑誌名 Nature Genetics	6. 最初と最後の頁 581~590
掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s41588-018-0067-2	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 Yasuda Jun, Katsuoka Fumiki, Danjoh Inaho, Kawai Yosuke, Kojima Kaname, Nagasaki Masao, Saito Sakae, Yamaguchi-Kabata Yumi, Tadaka Shu, Motoike Ikuko N., Kumada Kazuki, Sakurai-Yageta Mika, Tanabe Osamu, Fuse Nobuo, Tamiya Gen, Higasa Koichiro, Matsuda Fumihiko, Yasuda Nobufumi, Iwasaki Motoki, Sasaki Makoto et al.	4. 巻 19
2. 論文標題 Regional genetic differences among Japanese populations and performance of genotype imputation using whole-genome reference panel of the Tohoku Medical Megabank Project	5. 発行年 2018年
3. 雑誌名 BMC Genomics	6. 最初と最後の頁 551
掲載論文のDOI（デジタルオブジェクト識別子） 10.1186/s12864-018-4942-0	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 Kawaguchi S, Higasa K, Shimizu M, Yamada R, and Matsuda F.	4. 巻 38
2. 論文標題 HLA-HD: An accurate HLA typing algorithm for next-generation sequencing data.	5. 発行年 2017年
3. 雑誌名 Human Mutation	6. 最初と最後の頁 788-797
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/humu.23230	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 Kondo H, Uchio E, Kusaka S, Higasa K	4. 巻 39
2. 論文標題 Risk allele of FZD4 gene for familial exudative vitreoretinopathy.	5. 発行年 2018年
3. 雑誌名 Ophthalmic Genetics	6. 最初と最後の頁 405-406
掲載論文のDOI (デジタルオブジェクト識別子) 10.1080/13816810.2017.1401090	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 0件/うち国際学会 5件)

1. 発表者名 Ogata H, Higasa K, Kato T, Kageyama Y, Tahara H, Shimamoto A, Takekita Y, Bando H, Koshikawa Y, Sakai S, Nishida K, Nonen S, Kinoshita T, Kato M
2. 発表標題 THE RELATIONSHIP BETWEEN CIRCULATING MITOCHONDRIAL DNA AND MIRNA IN PATIENTS WITH MAJOR DEPRESSION
3. 学会等名 XXVIIth World Congress of Psychiatric Genetics (国際学会)
4. 発表年 2019年

1. 発表者名 Kato M, Ogata H, Higasa K, Tahara H, Shimamoto A, Tkekita Y, Bandou H, Kosikawa Y, Sakai S, Nishida K, Nonen S, Kinoshita T
2. 発表標題 MicroRNA Profiles as Predictor of Phenotypic Features of the Therapeutic Effect of Mirtazapine and SSRI in Mdd Patients
3. 学会等名 26th World Congress of Psychiatric Genetics (国際学会)
4. 発表年 2019年

1. 発表者名 Kawaguchi S, Tokumasu R, Numa S, Higasa K, Oishi A, Miyake M, Oishi M, Kamatani Y, Tsujikawa A, Takano A, Matsuda F
2. 発表標題 Development of AI technology to identify unknown causal genes and its integration in platform based on registry system.
3. 学会等名 International conference at Institut des Hautes Etudes Scientifiques (IHES) (国際学会)
4. 発表年 2018年

1. 発表者名 川口 修治、徳増 玲太郎、日笠 幸一郎、高野 敦司、松田 文彦
2. 発表標題 人工知能を用いたデータマイニングと全ゲノムデータ解析技術の融合によるメンデル遺伝性疾患の新規疾患遺伝子同定への試み
3. 学会等名 第63回日本人類遺伝学会 (国際学会)
4. 発表年 2018年

1. 発表者名 川口 修治、日笠 幸一郎、清水 正和、山田 亮、松田 文彦
2. 発表標題 次世代シーケンサーと国際データベースを用いた高効率・高精度なHLA遺伝子群の包括的タイピング技術の確立
3. 学会等名 第63回日本人類遺伝学会 (国際学会)
4. 発表年 2018年

〔図書〕 計2件

1. 著者名 川口修治、日笠幸一郎、山田亮、松田文彦	4. 発行年 2019年
2. 出版社 最新医学社	5. 総ページ数 6
3. 書名 最新医学 Vol.74(2)	

1. 著者名 Kawaguchi S, Higasa K, Yamada R, Matsuda F.	4. 発行年 2018年
2. 出版社 Humana Press	5. 総ページ数 9
3. 書名 HLA typing	

〔出願〕 計1件

産業財産権の名称 Pcr primer set for hla gene, and sequencing method using same	発明者 F.M. S.K. M. S. K. Higasa	権利者 同左
産業財産権の種類、番号 特許、US20190367983A1	出願年 2018年	国内・外国の別 外国

〔取得〕 計0件

〔その他〕

<p>てんかんの新しい発症機構の解明 繰り返し配列の異常伸長によっててんかんが生じることを発見 https://www.amed.go.jp/news/release_20180306-04.html</p> <p>Human Genetic Variation Database (HGVD) http://www.hgvd.genome.med.kyoto-u.ac.jp/</p>

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