

令和 4 年 6 月 10 日現在

機関番号：12602

研究種目：基盤研究(C) (一般)

研究期間：2019～2021

課題番号：19K07977

研究課題名(和文) ナノポアシーケンサーを用いた遺伝性神経疾患の解明

研究課題名(英文) Analysis of neurological genetic diseases using nanopore long read sequencer

研究代表者

三橋 里美 (Mitsubishi, Satomi)

東京医科歯科大学・難治疾患研究所・准教授

研究者番号：40466222

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：本研究では、20kb以上の塩基配列を連続して読むことのできるナノポア・ロングリード・シーケンサーを用いて、遺伝性疾患の原因となりうる単純反復配列やゲノム構造異常を検出する解析システムを構築し、比較的実験・解析コストが少なく疾患ゲノムを解析し、病的な変化を見出す解析系を構築した。この方法を用いて、従来の手法では未解決だった複数の疾患の原因を解き明かすことができた。ヒトゲノムの半分は反復する配列であり、まだその機能がわかっていない部分も多い。本研究によって、疾患の克服へとつながる疾患ゲノムの解明と、ヒトゲノムの未知の機能についての新たな知見が得られる一歩となったと考えている。

研究成果の学術的意義や社会的意義

従来型のシーケンサーは長い反復配列の検出に弱く、ヒトゲノムに100万以上存在すると言われる単純反復配列を網羅的に解析することは難しいと考えられており、ロングリードシーケンサーを用いたゲノム解析の開発に期待がもたれていた。本研究では、実際の疾患患者の解析にロングリードシーケンサーを用いて、未解決課題を解決することができるというproof of conceptを示すことができた。さらに、誰もが難治性疾患の診断や研究に使うことができるように、コストを抑えたターゲットシーケンスによる解析系を立ち上げたことは意義が大きいと考えている。

研究成果の概要(英文)：In this study, we developed an analysis system to detect tandem repeats and genomic structural abnormalities that may cause genetic diseases using a nanopore long-read sequencer. We analyzed actual patients' genomes with relatively low analysis cost and labor to find pathological alterations. Using this method, we were able to elucidate the causes of multiple diseases that were unresolved by conventional methods. Genomic functions are not yet understood in many parts of the genome; most of which are repetitive regions. We believe that our attempt to unveil the disease genome will lead to overcoming diseases and will be a step toward gaining new insights about the unknown functions of the human genome.

研究分野：神経内科学

キーワード：ナノポアシーケンサー 遺伝性神経疾患 ロングリードシーケンサー 単純反復配列

1. 研究開始当初の背景

遺伝性疾患の原因変異探索において、この10年間、大量の短い塩基配列を並列に大量に解析する次世代シーケンサーによる網羅的な遺伝子配列解析が行われてきた。しかし、この方法は、CTGなど数塩基が非常に長く繰り返す配列(単純反復配列)や、LINE、SINEなど同じ塩基配列がゲノム上に繰り返し存在する領域(散在性反復配列)の解析が、時として難しい場合がある。高度に類似した配列は、変異率が一塩基置換よりも高いことがあり、様々なゲノム構造の変化を引き起こす。ヒトゲノムには、単純反復配列のコピー数変化、欠失、重複、挿入、転座、逆位などの構造多型が数多く存在し、こういった構造変化が重篤な疾患を引き起こすことが知られている。これらの反復配列の解析には、1万塩基以上を一つづきに読むことのできる、ロングリードシーケンサーによるゲノム解析が有利であると考えた。研究開始当初、ロングリードシーケンサーはエラー率も高く、解析のための基盤となるインフォマティクス技術や、プラットフォームが整備されていなかったため、これらのゲノム構造異常をロングリードから解析する2つのツールの構築とその解析システムの実用化を計画した。この手法によって、従来手法で見落とされてきた、単純反復配列の伸長による疾患や、複雑なゲノム構造異常の詳細が解き明かされるのではないかと考え、研究を開始した。

2. 研究の目的

本研究では、ナノポアシーケンサーを用いて、遺伝性疾患の原因となりうる単純反復配列やゲノム構造異常を検出する解析システムを構築し、そのヒト疾患における意義や成り立ちを解明することを目指す。

3. 研究の方法

単純反復配列の変化をゲノム網羅的に解析するツール tandem-genotypes と、新しく開発したゲノム構造異常を検出するツール dnarrange を使い、従来手法で原因が不明だった神経疾患などの、遺伝性疾患の全ゲノムシーケンズデータの解析を行う。さらに、lamassemble というロングリードを多重アライメントするツールの開発も行い、伸長リピートや構造異常のコンセンサス配列を作ることで、正確な切断点やリピート内部構造を決定する。これらのツールには、患者で特に伸長している反復配列に優先順位をつけることや、他のコントロールにはみられないまれな構造多型を絞り込むなどの機能があるため、疾患原因となりうる稀な変化を効率的に検出することが可能である。

4. 研究成果

単純反復配列については、神経核内封入体病の原因が NOTCH2NL の 5' UTR にある GGC リピートの伸長であることを見出していたが(Sone & Mitsuhashi & Fujita et al. 2019)、その後も複数の患者を同定し(Okubo et al. 2019)、さらにロングリードシーケンスによるエピゲノム解析をおこなった(Fukuda et al. 2021)。また、これまで原因不明であった小脳性運動失調-ニューロパチー-前庭反射消失症候群(CANVAS)の一例に、疾患原因となる RFC1 のイントロン領域に両アレルのリピート伸長を見出した (Nakamura et al. 2020)。さらに、Iamassemble というロングリードを多重アライメントする手法により、ロングリードのコンセンサス配列を作り単純反復配列内部の構造や構造多型を明らかにすることができることが可能となった(Frith et al. 2021)。これは、リピート内の構造が表現系と関わることや、精度高く染色体の切断点を同定できることから、疾患メカニズムを考える上でも非常に有用であった。

さらに、過去にショートリードシーケンサーやマイクロアレイでは詳細が同定できなかった 5 例の染色体異常患者について、ナノポアシーケンサーで全ゲノムシーケンスを行い、dnarrange 解析によりゲノム構造を明らかにした(Lei et al. 2020, Mitsuhashi & Ohori et al. 2021)。そのうち 2 例では多数の複雑なゲノム構造異常を見出し、染色体破碎現象が示唆された。同時に開発した dnarrange-link を用いて、複雑なゲノム構造を組み立て、リファレンスゲノムと比較し、ゲノム構造の変化を再構築することにより、染色体破碎現象を可視化することに成功した。また、dnarrange では、同じ構造異常を持たないロングリードシーケンス データをコントロールにすることにより、数千あった構造多型から、患者特有の構造多型を 100 以下に絞り込むことができ、疾患原因となるまれな多型を効率的に見出すことが重要であることがわかった。このような絞り込みはロングリードの構造多型データでは行われたことがなく、新規性が高い。

さらに、疾患原因となるリピートは健常においても多型性があるという結果を得た(Mitsuhashi et al. 2021)。これは多型性のあるリピートに注目して新たな疾患原因を見出すことができるかもしれないということを示唆しており、過去にゲノムワイド関連解析や疾患家系の連鎖解析などで見つかった疾患関連ローカスの近くにある多型性のあるリピートが原因となっている可能性について、神経疾患をターゲットに新たなリピートの原因を絞り、複数の箇所を同時にターゲットシーケンスする方法を確立した。

また、遺伝性疾患の中には、散在性反復配列からの転写産物の発現が上昇しているものがあり、このメカニズムを明らかにするため、ナノポアシーケンサーによる direct RNA シーケンスを用いて、この疾患で特徴的に発現する新規のトランスクリプトを多数同定した(Mitsuhashi et al. 2021)。現在、これらの反復配列の機能解析を進めている。

遺伝性疾患の原因は、従来のショートリードシーケンサーで明らかになっていないものが半数以上存在する。ロングリードシーケンサーの実用化によって、これまでの技術では困難であった領域の解析が期待されているが、臨床遺伝学的に実用的なツールがほとんどない。本研究では、新しく開発した tandem-genotypes と dnarrange という二つのアプローチを用いることで、遺伝性疾患のロングリード全ゲノム解析結果から、疾患原因を見出すことが可能になることを示しており、さらなる実用化へむけてこれらのツール開発を行っていきたいと考えている。

5. 主な発表論文等

〔雑誌論文〕 計9件（うち査読付論文 8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 Mitsuhashi Satomi, Nakagawa So, Sasaki-Honda Mitsuru, Sakurai Hidetoshi, Frith Martin C, Mitsuhashi Hiroaki	4. 巻 30
2. 論文標題 Nanopore direct RNA sequencing detects DUX4-activated repeats and isoforms in human muscle cells	5. 発行年 2021年
3. 雑誌名 Human Molecular Genetics	6. 最初と最後の頁 552-563
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/hmg/ddab063	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 Frith Martin C., Mitsuhashi Satomi, Katoh Kazutaka	4. 巻 2231
2. 論文標題 Iamassemble: Multiple Alignment and Consensus Sequence of Long Reads	5. 発行年 2020年
3. 雑誌名 Methods Mol Biol.	6. 最初と最後の頁 135-145
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-1-0716-1036-7_9	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 Mitsuhashi Satomi, Ohori Sachiko, Katoh Kazutaka, Frith Martin C., Matsumoto Naomichi	4. 巻 12
2. 論文標題 A pipeline for complete characterization of complex germline rearrangements from long DNA reads	5. 発行年 2020年
3. 雑誌名 Genome Medicine	6. 最初と最後の頁 67
掲載論文のDOI（デジタルオブジェクト識別子） 10.1186/s13073-020-00762-1	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 Lei Ming, Liang Desheng, Yang Yifeng, Mitsuhashi Satomi, Katoh Kazutaka, Miyake Noriko, Frith Martin C., Wu Lingqian, Matsumoto Naomichi	4. 巻 65
2. 論文標題 Long-read DNA sequencing fully characterized chromothripsis in a patient with Langer?Giedion syndrome and Cornelia de Lange syndrome-4	5. 発行年 2020年
3. 雑誌名 Journal of Human Genetics	6. 最初と最後の頁 667-674
掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s10038-020-0754-6	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 Mitsuhashi Satomi, Frith Martin C., Matsumoto Naomichi	4. 巻 14
2. 論文標題 Genome-wide survey of tandem repeats by nanopore sequencing shows that disease-associated repeats are more polymorphic in the general population	5. 発行年 2021年
3. 雑誌名 BMC Medical Genomics	6. 最初と最後の頁 17
掲載論文のDOI (デジタルオブジェクト識別子) 10.1186/s12920-020-00853-3	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 Satomi Mitsuhashi, Naomichi Matsumoto	4. 巻 65
2. 論文標題 Long-read sequencing for rare human genetic diseases	5. 発行年 2020年
3. 雑誌名 Journal of Human Genetics	6. 最初と最後の頁 11-19
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s10038-019-0671-8	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 Jun Sone, Satomi Mitsuhashi, Atsushi Fujita et al.	4. 巻 51
2. 論文標題 Long-read Sequencing Identifies GGC Repeat Expansions in NOTCH2NLC Associated With Neuronal Intranuclear Inclusion Disease	5. 発行年 2019年
3. 雑誌名 Nature Genetics	6. 最初と最後の頁 1215-1221
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41588-019-0459-y	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 Haruko Nakamura, Hiroshi Doi, Satomi Mitsuhashi, Satoko Miyatake, Kazutaka Katoh, Martin C Frith, Tetsuya Asano, Yosuke Kudo, Takuya Ikeda, Shun Kubota, Misako Kunii, Yu Kitazawa, Mikiko Tada, Mitsuo Okamoto, Hideto Joki, Hideyuki Takeuchi, Naomichi Matsumoto, Fumiaki Tanaka	4. 巻 65
2. 論文標題 Long-read Sequencing Identifies the Pathogenic Nucleotide Repeat Expansion in RFC1 in a Japanese Case of CANVAS	5. 発行年 2020年
3. 雑誌名 Journal of Human Genetics	6. 最初と最後の頁 475-480
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s10038-020-0733-y	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 Ming Lei, Desheng Liang, Yifeng Yang, Satomi Mitsuhashi, Kazutaka Katoh, Noriko Miyake, Martin C Frith, Lingqian Wu, Naomichi Matsumoto	4. 巻 65
2. 論文標題 Long-read DNA Sequencing Fully Characterized Chromothripsis in a Patient With Langer-Giedion Syndrome and Cornelia De Lange syndrome-4	5. 発行年 2020年
3. 雑誌名 Journal of Human Genetics	6. 最初と最後の頁 667-674
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s10038-020-0754-6	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3件 (うち招待講演 0件 / うち国際学会 2件)

1. 発表者名 Satomi Mitsuhashi
2. 発表標題 Finding complex pathogenic variants using nanopore long reads
3. 学会等名 American Society of Human Genetics 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 三橋里美
2. 発表標題 ロングリードシーケンスによるタンデムリピート領域の解析
3. 学会等名 第60回日本神経学会学術大会
4. 発表年 2019年

1. 発表者名 三橋里美
2. 発表標題 Finding complex mutations using long reads
3. 学会等名 第13回国際ゲノム会議 (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔出願〕 計1件

産業財産権の名称 神経核内封入体病患者の検出方法	発明者 祖父江元、曾根淳、 松本直通、三橋里 美、藤田京志	権利者 同左
産業財産権の種類、番号 特許、1778	出願年 2019年	国内・外国の別 外国

〔取得〕 計0件

〔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	Frith Martin (Frith Martin) (40462832)	東京大学・大学院新領域創成科学研究科・教授 (12601)	

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