

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 5 月 31 日現在

機関番号：22701

研究種目：基盤研究（A）

研究期間：2009～2011

課題番号：21249024

研究課題名（和文）高速シーケンサー解析を効率化するゲノム領域選択技術の開発研究

研究課題名（英文）Development of genome-partitioning technologies for next generation sequencing

研究代表者

松本 直通（MATSUMOTO NAOMICHI）

横浜市立大学・医学研究科・教授

研究者番号：80325638

研究成果の概要（和文）：PCR によるゲノム領域をカバーするゲノム領域特異的増幅法、150-200mer のビオチンラベル cRNA を領域特異的プローブとして溶液中でハイブリさせる液層ハイブリ法を用いて次世代シーケンサー解析を行った。PCR システムの次世代シーケンス解析でシーケンス正確性と検出度を確認した。液相ハイブリ法は、対象領域のほぼ90%の領域の検証が可能であり網羅的スクリーニングの系として有用であった。現在後者のシステムを採用し疾患遺伝子を行っている。

研究成果の概要（英文）：The advent and frequent update of next generation sequencers (NGSs) can attain the appropriate accuracy for mutation analysis and push disease-related genome analysis into the new stages. We now use Illumina Genome Analyzer (GA) IIx and HiSeq2000 which can produce as much as 60-Gb and 600-Gb sequences in one run, respectively. To focus on genes, we utilized exon capture methods such as SureSelect (Agilent). The current NGS protocol uses 100-108-bp pair-end reads and usually produces 8-9 Gb sequences (per one sample) could be enough for analysis of the whole exome: 90% of exome bait regions are covered by 8-10 reads or more. Sequences are aligned using MAQ, BWA, Novoalign and commercial-based NextGENe software all of which are able to extract nucleotide changes and small insertions/deletions. The most critical step is the priority scheme selecting variants. We have been successful in addressing culprit mutations in several diseases.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	19,000,000	5,700,000	24,700,000
2010年度	9,200,000	2,760,000	11,960,000
2011年度	9,200,000	2,760,000	11,960,000
年度			
年度			
総計	37,400,000	11,220,000	48,620,000

研究分野：医歯薬学

科研費の分科・細目：基礎医学・人類遺伝学

キーワード：分子遺伝学

1. 研究開始当初の背景

一般的なサンガー法を用いたヒト1名のディプロイドゲノム分(60億塩基対)の完全な塩基配列決定には、約1000万ドル(12億円)のコストが必要と試算される。次世代シーケンサーのロシユ 454、イルミナ Solexa、ABI SOLID を用いれば、それぞれ、0.05 Gb(ロシユ)、1 Gb(イルミナ)、3 Gb(ABI)規模のシーケンス解読が1回で可能である。現行のサンガー法の100-1000倍の処理能力と1/100-1/1000のコストが実現される(イルミナ Solexa での試算)。本研究は高速シーケンサーを用いてヒト疾患ゲノム解析を行うために鍵となると考えられるゲノム領域を選択する技術を用いて、この技術と高速シーケンサーを組み合わせる効果的なヒト疾患責任遺伝子単離・同定法を確立することを目的に提案された。

2. 研究の目的

所有する高速シーケンサーはイルミナ社・GAIIxである。2008年のカタログスペックでは、1回のシングルランで36bpのシーケンス長で5千万リードが可能で産出シーケンスは1.5Gbである。Solexaは市販されている次世代シーケンサーの中で最も普及し、反応系がシンプルでデータ産出が容易さにおいて優位性がある。ヒト遺伝子領域の変異・SNP検出能力においては、高いシーケンス産出力による対象領域の高カバー率を考慮すると次世代機の中で遺伝子変異解析において優位性がある。さらに解析領域のゲノムサイズを20-30Mb程度に絞り込む(選択する)ことで、1ランで20カバレッジ以上の解析は容易に可能となりシーケンスエラーは、1Mbに1-2個程度と変異解析において大きな障害にならない解析系となる。ヒト遺伝子領域は全ゲノムの約1%であり約30Mbがエクソン領域となるため全遺伝子エクソン領域の解析も1ランの解析対象レベル(50カバレッジ)である。よっ

て解析領域を効率よく選択・回収できれば1例のエクソン領域のdeep sequencingが1ランで可能となり、遺伝子の突然変異によるメンデル遺伝病などの解析には最適な解析系となる。ゲノム領域選択技術として①Long-PCRによる増幅法、②アレーを用いたシーケンスキャプチャー法(下図左)、③液層ゲノム分画法(下図右)が挙げられる。①はLong-PCRで10-25KbのPCR増幅を行いこの産物によって解析対象領域をカバーする方法である。現実ではあるがPCRエラーと多数のPCRプライマーを含む試薬コストが問題である。例えば1Mbの領域は20KbのLong-PCRが50セットでカバー(20x50=1000Kb)される。②はアレーに固着した領域特異的DNAプローブで、③は液層に浮遊させた領域特異的cRNAプローブで、ゲノム解析領域を選択するもので、最大30Mb程度のゲノム領域を回収することが可能でありこれを次世代機でシーケンス解析する。

解析対象疾患は、単一遺伝子病(メンデル遺伝病)である。申請者らはこれまでの研究過程で責任遺伝子が未解決の複数のメンデル遺伝性疾患の検体の集積がある。遺伝子責任領域マッピングが完了しているが遺伝子未同定の脊髄小脳変性症、Waardenburg 無眼球症候群や新型エーラスダンロス症候群、あるいは遺伝子への手がかりがない歌舞伎メーキャップ症候群、Coffin-Siris症候群、Aicardi症候群などである。これらの症例は、全て保有する高密度マイクロアレー解析を行い病的な染色体微細欠失・重複を有さないことを確認済みで、点変異による疾患発症が極めて濃厚な症例群を解析対象とする。

3. 研究の方法

比較検討するゲノム領域選択技術は以下の3種類である。

(1) Long-PCRによる領域増幅法

近年のPCR用酵素の進歩でゲノムDNAを鋳型にしたPCRですら10-25Kbサイズの増幅が可能になってきている。この技術を用い

れば 1 Mb の領域は約 50 個の PCR で被覆できる。これらの PCR プロダクトを次世代シーケンサーで解析すれば 1 ランで解読終了となる。PCR であるため検出された塩基変化は PCR エラーと区別する必要があるが、それぞれの変化は個別に症例とその両親で通常の PCR シーケンスによるトリオ検証等で解釈は容易である。

(2) アレーを基盤とするシーケンスキャプチャー法

60mer のオリゴ DNA からなるシーケンスキャプチャー用マイクロアレーを作成する。高密度オリゴ DNA アレーを作成すれば最大 30 Mb 程度の領域をカバーすることが可能である。解析対象ゲノム DNA を 500 bp 程度に断片化し両端に linker-primer を付けアレーにハイブリ、そして領域選択を行う。選択されたゲノム断片はプライマーで増幅後、次世代シーケンサー解析を行う。アレーにハイブリし回収された DNA が、ヒトゲノムの両アレルを代表しているか、対象領域をカバーしているかの評価が重要である。NimbleGen 社や Agilent 社のプラットフォームで設計が可能である。

(3) 液層ゲノム分画法

ビオチン標識化した 150~200mer の cRNA をゲノム領域選択プローブとする。このプローブを 5 万~6 万種類作成したものを 1 チューブ内に収納し 500 bp 程度に断片化した疾患ゲノムをチューブ内の溶液中でハイブリダイゼーションさせる。ハイブリさせたプローブをストレプトアビジンを固着したマグネットビーズで回収し、プローブ RNA を消化しハイブリダイズした領域特異的 DNA を収集する。cRNA プローブが 150-200mer と長く特異性が高いことと cRNA プローブ量が液層で濃度が高く設定できるため領域特異的 DNA 回収効率が高いことが期待される。回収したゲノム DNA を次世代シーケンサーで解析する。液層分画プラットフォームは Agilent 社で設計が可能である。

(4) 最適プラットフォームの比較検討と選択

以上のゲノム領域選択技術を同じ症例で解析しそれぞれの解析データを比較検討し、

コスト並びにゲノム選択効率から最良のプラットフォームを決定する。単一のプラットフォームで網羅的な解析が難しい場合、それぞれの解析を組み合わせて補完しながら候補領域の解析を行う必要がある。

(5) 解析対象

メンデル遺伝性疾患で責任遺伝子が未同定の複数の疾患が対象である。また正常表現型両親に生まれた典型孤発例が最適の対象となる。即ち、両親には観察されない遺伝子異常が症例にのみ観察され新生突然変異 (fresh mutation) による発症が明らかにできる症例が理想的である。遺伝子責任領域マッピングが完了しているが遺伝子未同定の脊髄小脳変性症や新型エーラスダンロス症候群、あるいは遺伝子への手がかりがない歌舞伎メーキャップ症候群、Coffin-Siris 症候群、Aicardi 症候群などを解析したいと考えている。既にそれぞれの疾患で 10~40 例程度の蓄積があり候補遺伝子の変異を見つけた場合直ぐに普遍化できる、即ちその他の症例で検証できる状況にある。

4. 研究成果

高速シーケンサーを用いてヒト疾患ゲノム解析を行うために鍵となるゲノム領域を選択する技術を用いて、この技術と高速シーケンサーを組み合わせた効果的なヒト疾患責任遺伝子単離・同定法を確立することを目的として研究を進めた。PCR によるゲノム領域をカバーするゲノム領域特異的増幅法、150-200mer のビオチンラベル cRNA を領域特異的プローブとして溶液中でハイブリさせる液層ハイブリ法、あるいはマイクロアレーを基盤とした固層ハイブリ法を用いて遺伝性要因の高いと考えられる複数の疾患を対象に次世代シーケンサー解析を行った。PCR システムの次世代シーケンス解析とリシーケンシングアレーを用いた検証でほぼ同等のシーケンス正確性と検出度を確認することができた。このうち液相ハイブリ法と固層ハイブリ法は、対象領域のほぼ 90% の領域の検証が可能であり網羅的スクリーニングの系として有用であった。産出したゲノムシーケンス情報は、MAQ、BWA、

Novoalign 等の各種 aligner でマッピングを施行し、現行のデータベース上のヒトゲノム参照シーケンスと SNP データベース情報を参考に変異候補となるバリエントを抽出、そのリストを作製した。個々のバリエントについて症例とその両親のゲノムでの確認を行い、いくつかの疾患で新生突然変異 (fresh mutation) を同定した。これまでのところ Novoalign と GATK を用いてバリエントを抽出し ANNOVA を用いて意義付けを行うシステムを採用作り上げ疾患遺伝子の特定に効果を上げている。脊髄小脳変性症や歌舞伎メーキャップ症候群、Coffin-Siris 症候群で遺伝子特定に至った。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 10 件)

- ① Saitsu H, Tohyama J, Kumada T, Egawa K, Hamada K, Okada I, Mizuguchi T, Osaka H, Miyata R, Furukawa T, Haginoya K, Hoshino H, Goto T, Hachiya Y, Yamagata T, Saitoh S, Nagai T, Nishiyama K, Nishimura A, Miyake N, Komada M, Hayashi K, Hirai S, Ogata K, Kato M, Fukuda A, *Matsumoto N. Dominant negative mutations in α -II spectrin cause early onset West syndrome with severe cerebral hypomyelination, spastic quadriplegia, and developmental delay. *Am J Hum Genet* 86(6):881-889, 2010.
- ② Osaka H, Yamamoto R, Hamanoue H, Nezu A, Sasaki M, Saitsu H, Kurosawa K, Shimbo H, Matsumoto N, Inoue K. Disrupted SOX10 regulation of GJC2 transcription causes Pelizaeus-Merzbacher-Like Disease. *Ann Neurol* 68(2): 250-254, 2010
- ③ Miyake N, Kosho T, Mizumoto S, Furuichi T, Hatamochi A, Nagashima Y, Arai E, Takahashi K, Kawamura R, Wakui K, Takahashi J, Kato H, Yasui H, Ishida T, Ohashi H, Nishimura G, Shiina M, Saitsu H, Tsurusaki T, Doi H, Fukushima Y, Ikegawa S, Yamada S, Sugahara K, Matsumoto N. Loss of decorin dermatan sulfate impairing collagen bundle formation in a new type of Ehlers-Danlos syndrome. *Hum Mut* 31(8): 966-974, 2010
- ④ Ng S, Bigham A, Buckingham K, Hannibal M, McMillin M, Gildersleeve H, Beck A, Tabor H, Cooper G, Mefford H, Lee C, Turner E, Smith J, Rieder M, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson D, *Bamshad M, *Shendure J. Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome. *Nat Genet* 42(9): 790-793, 2010.
- ⑤ Okada I[#], Hamanoue H[#], ([#] denotes equal contribution) Terada K, Tohma T, Megarbane A, Chouery E, Abou-Ghoch J, Jalkh N, Cogulu O, Ozkinay F, Horie K, Takeda J, Furuichi T, Ikegawa S, Nishiyama K, Miyatake S, Nishimura A, Mizuguchi T, Niikawa N, Hirahara F, Kaname T, Yoshiura K-i, Tsurusaki Y, Doi H, Miyake N, Furukawa T, *Matsumoto N, *Saitsu H. *SMO1* is essential for ocular and limb development in humans and mice. *Am J Hum*
- ⑥ Tsurusaki Y, Osaka H, Hamanoue H, Shimbo H, Tsuji M, Doi H, Saitsu H, *Matsumoto N, *Miyake N. Rapid detection of a mutation causing X-linked leukodystrophy by exome sequencing. *J Med Genet* 48 (9): 606-609, 2011.
- ⑦ Saitsu H, Matsumoto N. Genetic commentary: *De novo* mutations in epilepsy. *Dev Med Child Neurol* 53 (9):806-807, 2011.
- ⑧ Saitsu H, Osaka H, Sasaki M, Takanashi J, Hamada K, Yamashita A, Shiina M, Kondo Y, Nishiyama K, Tsurusaki Y, Miyake N, Doi H, Ogata K, Inoue K, *Matsumoto N. Mutations in *POLR3A* and *POLR3B* encoding RNA polymerase III

subunits cause an autosomal recessive hypomyelinating leukoencephalopathy. *Am J Hum Genet* 90 (1):86-90, 2011.

- ⑨ Tsurusaki Y, Okamoto N, Suzuki Y, Doi H, Saitsu H, Miyake N, Matsumoto N. Exome sequencing of two patients in a family with atypical X-linked leukodystrophy. *Clin Genet* 80 (2): 161-166, 2011.
- ⑩ Doi H, Yoshida K, T Yasuda, Fukuda M, Fukuda Y, Morita H, Ikeda S-i, Kato R, Tsurusaki Y, Miyake N, Saitsu H, Sakai H, Miyatake S, Shiina M, Nukina N, Koyano S, Tsuji S, Kuroiwa Y, *Matsumoto N. Exome sequencing reveals a homozygous *SYT14* mutation in adult-onset autosomal recessive spinocerebellar ataxia with psychomotor retardation. *Am J Hum Genet* 89(2):320-327, 2011.

[学会発表] (計 16 件)

- ① 第 44 回遺伝医学研究会 (東京女子医科大学) (6 月 5 日・東京女子医科大学・早稲田大学連携生命科学教育施設) 松本直通「年齢依存性てんかんの最近の話題」(招聘講演)
- ② 第 44 回日本小児腎臓病学会学術集会 (6 月 26 日・一橋記念講堂・東京) 松本直通「疾患ゲノム解析の新戦略」(特別講演)
- ③ 第 51 回日本小児神経学会関東地方会 (9 月 19 日横浜市開港記念会館) 松本直通「年齢依存性てんかん性脳症の最近の話題」(特別講演)
- ④ The international symposium in the 9th Annual Meeting of the East Asian Union of Human Genetic Society (Nov 19, 2009, Soel, Korea) Naomichi Matsumoto: *STXBPI* mutations in severe infantile epilepsies with suppression-burst pattern (oral presentation).
- ⑤ Asin Cytogenetics Community Workshop (by Affymetrix Inc.) Naomichi Matsumoto. Evaluation of Affymetrix® Cytogenetics Whole-Genome Array Using Clinical Sample. (Oct 24, 2009 at Honolulu, Hawaii, HI) (invited lecture)
- ⑥ 第 113 回日本小児科学会学術集会 (4 月 23 日於岩手県民情報交流センター、盛岡)・シンポジウム「先天性疾患における最近の進歩：病態解明から遺伝子診断へ」・松本直通「染色体異常からの疾患遺伝子探索」(シンポジスト)
- ⑦ Joint Egyptian-Japanese Scientific Workshop, “A new era of genetic diseases” (organized by Ghada M.H. Abdel Salam and Naomichi Matsumoto) (National Research Center, Cairo, Egypt, Oct 3-4, 2010) Naomichi Matsumoto “Microarray technologies: Highways to genomic aberrations”
- ⑧ Joint Egyptian-Japanese Scientific Workshop, “A new era of genetic diseases” (organized by Ghada M.H. Abdel Salam and Naomichi Matsumoto) (National Research Center, Cairo, Egypt, Oct 3-4, 2010) Naomichi Matsumoto “Isolation of the gene responsible for a new type of Ehlers-Danlos syndrome”
- ⑨ Joint Egyptian-Japanese Scientific Workshop, “A new era of genetic diseases” (organized by Ghada M.H. Abdel Salam and Naomichi Matsumoto) (National Research Center, Cairo, Egypt, Oct 3-4, 2010) Naomichi Matsumoto “Haploinsufficiency of *STXBPI* causes Ohtahara syndrome”
- ⑩ The 4th Asian Chromosome Colloquium (Beijing, China, Oct 11-14) Naomichi Matsumoto (invited speaker) “Identification of two epilepsy-related genes from a 2.25-Mb deletion in one patient.”
- ⑪ 日本人類遺伝学会第 55 回大会 (大宮、10 月 30 日) 松本直通「疾患ゲノム解析：遺伝性疾患のエクソーム解析」(次世代シーケンサーを用いたヒト (疾患) ゲノム解析の現状セッション・シンポジスト・座長)

- ⑫ 第22回 NIH 金曜会 (National Institute of Health, Bethesda, MD 11月5日)
Naomichi Matsumoto (Invited lecture)
「Identification of two genes responsible for age-dependent epileptic encephalopathy」
- ⑬ Naomichi Matsumoto “Identification of two epilepsy-related genes from a 2.25-Mb microdeletion in one patient”
(Invited lecture at Department of Human Genetics, Leiden University, Leiden, The Netherland, May 26, 2011)
- ⑭ 第18回遺伝性疾患に関する出生前診断研究会「次世代シーケンサーを用いた疾患ゲノム解析の現状」(佐賀大学医学部・佐賀10月1日)
- ⑮ 日本人類遺伝学会第56回大会「ヒト遺伝性疾患の原因解明を目指して」学会賞受賞講演(於・幕張メッセ11月11日)
- ⑯ 日本人類遺伝学会第56回大会「次世代シーケンサーを用いたヒト疾患ゲノム解析法」(シンポジスト)シンポジウム11(超高速シーケンサーによる疾患ゲノム解析)(於・幕張メッセ11月12日)

[産業財産権]

○出願状況(計7件)

名称:び慢性大脳白質形成不全症の検出方法

発明者:才津浩智/松本直通

権利者:横浜市立大学

種類:特願

番号:2011-226488

出願年月日:平成23年10月14日

国内外の別:国内

他6件

○取得状況(計0件)

[その他]

ホームページ等

http://www.yokohama-cu.ac.jp/amedrc/res/matsumoto2011_8.html

<http://www.yokohama-cu.ac.jp/amedrc/res/>

[matsumoto2011_10.html](http://www.yokohama-cu.ac.jp/amedrc/res/matsumoto2011_10.html)

http://www.yokohama-cu.ac.jp/amedrc/res/saitsu2011_12.html

http://www.yokohama-cu.ac.jp/amedrc/res/matsumoto2012_03.html

http://www.yokohama-cu.ac.jp/amedrc/res/matsumoto2012_0319.html

6. 研究組織

(1) 研究代表者

松本 直通 (MATSUMOTO NAOMICHI)

横浜市立大学・医学研究科・教授

研究者番号:80325638

(2) 研究分担者

三宅 紀子 (MIYAKE NORIKO)

横浜市立大学・医学部・准教授

研究者番号:40523494

(3) 連携研究者

なし