

平成 28 年 4 月 26 日現在

機関番号：13802

研究種目：挑戦的萌芽研究

研究期間：2014～2015

課題番号：26670505

研究課題名(和文) ターゲットRNAキャプチャーを用いた疾患責任遺伝子の網羅的RNAシーケンス解析

研究課題名(英文) Comprehensive RNA analysis of disease causing genes by target RNA capture

研究代表者

才津 浩智 (Saito, Hirotomo)

浜松医科大学・医学部・教授

研究者番号：40402838

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：スプライス異常が既知の患者検体7検体と、常染色体劣性遺伝形式が想定される遺伝子の片アレル変異のみを有している1検体の計8検体(リンパ芽球)について、24遺伝子をターゲットにしたRNAキャプチャーを施行し、ライナーの性能を比較した。Novoalignはエクソンスキッピングの検出感度がよい傾向にあったが、イントロン保持のスプライス異常は検出できなかった。TopHatは両方の異常が検出可能であり、ライナーとして有用であると考えられた。また、遺伝子毎の発現量に極めて大きな差が認められ、検査不能の遺伝子が5遺伝子あり、病気の場となる組織のRNAを得ることが重要だと考えられた。

研究成果の概要(英文)：Lymphoblastoid cells derived from seven patients with mutations causing abnormal splicing and a patient with a possible mutation only in one allele were examined by target RNA capture against 24 genes. Enriched libraries were sequenced in one lane of HiSeq 2500 sequencer, and reads were aligned by Novoalign to transcriptome or by TopHat to reference genome. Although Novoalign was able to align more reads with exon skipping, it failed to align reads with intron retention. On the other hand, TopHat was able to align both reads with exon skipping and intron retention, suggesting that TopHat is first choice for alignment. In lymphoblastoid cells, five out of 24 genes are not well sequenced, indicating that even in target capture system, tissues responsible for disease are required for optimized analysis.

研究分野：分子遺伝学

キーワード：スプライス変異 ターゲットキャプチャー 次世代シーケンス

1. 研究開始当初の背景

一般的に、常染色体劣性遺伝病は症状が重篤であり、責任遺伝子と表現型の相関が明らかであることが多い。なかには、遺伝子治療（芳香族アミノ酸脱炭酸酵素欠損症）や骨髄移植（クラベ病）といった治療が有効な疾患もあるが、治療自体が挑戦的あるいは患者の負担が大きく、治療に踏み切るためには遺伝子診断を行うことが重要となる。近年の次世代シーケンサーの登場とエクソーム解析の開発により、網羅的な遺伝子解析が可能となり、既知の責任遺伝子の変異が同定される症例が飛躍的に増加した。しかし、常染色体劣性遺伝病では、2 つある遺伝子座の一方にしか変異が見つからないという症例が少なからず存在し、エクソーム解析で解析不能なイントロンの変異によるスプライシング異常の関与等が疑われていた。そのため、このような症例における、スプライシング異常をきたすような変異を同定する解析方法の確立が期待される。1つの方法として、異常の結果、すなわち RNA のスプライシング異常の同定を先に行い、異常が同定された場合は該当するゲノム上の変異を丹念に探していくという方法が考えられる。最近開発された、特定の遺伝子群の mRNA 配列をキャプチャー後にシーケンスを行う、ターゲット RNA シーケンス法はこのスプライシング異常を高感度に検出できる可能性を秘めている。

2. 研究の目的

本研究では、RNA キャプチャーを用いた疾患責任遺伝子の網羅的 RNA シーケンスを通じて、スプライシング異常の高感度検出が可能な、画期的な遺伝子変異解析系の確立を目指した。また、逆転写 PCR 産物を次世代シーケンスした際の、最適なアライメントの方法の検討も行った。

3. 研究の方法

(1) スプライシング異常が既知の患者検体（リンパ芽球）7 検体と、常染色体劣性遺伝形式が想定される遺伝子に片アレルのみ変異を有している 1 検体（リンパ芽球）の計 8 検体において、24 遺伝子をターゲットにした RNA キャプチャーを施行し、得られたキャプチャーライブラリを MiSeq あるいは HiSeq2500 の 1 レーンでシーケンスした。

(2) 特定の遺伝子変異がもたらすスプライシング異常を網羅的に解析するために、逆転写 PCR 産物にアダプターを付加して、MiSeq でシーケンスし、異常スプライシング産物の同定およびその発現量の定量化を行った。

4. 研究成果

(1) まず MiSeq でシーケンスを行い、1 検体当たり 150 Mb のインプット量を得た。このデータを用いて、TopHat と novoalign の 2 つのアライナーを用いて検討し、エクソン

スキッピングとイントロン保持のパターンのスプライシング異常の検出について比較検討を行った。Novoalign の方がエクソンスキッピングの検出感度がよい傾向がえられたが、反対にイントロン保持のスプライシング異常は検出できなかった。これは、転写産物配列に対してアライメントしているためと考えられる。一方、TopHat はイントロン保持のスプライシング異常、エクソンスキッピング異常の両方の異常が検出可能であった。これらの結果から、TopHat でのアライメントが第一選択となると考えられた。

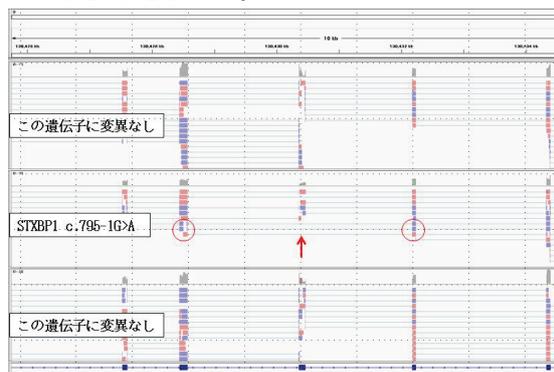


図 1 : MiSeq でシーケンスしたリードを TopHat でアライメントした。真ん中がスプライシング部位に変異を有数する症例。エクソンのスキップ (矢印) が認められる。

次に、より厚みを持ってシーケンスすることを目的として HiSeq2500 の 1 レーンで 8 検体をシーケンスした。TopHat でアライメントした場合の遺伝子毎の平均のリードの厚みの表を下に示す (4 例のみ示す)。

	Patient 1	Patient 2	Patient 3	Patient 4
PARS2	35451	106661	55901	89845
PLB1	46	130	88	122
GALNT13	0	5	13	6
QARS	470329	450996	435778	502306
CNPY3	1605575	788004	1342329	962337
FAM126A	78066	205256	95258	161047
PNPLA8	28998	185811	28545	98505
B4GALT1	20338	423427	25075	249237
STXBP1	638	565	397	481
POLR3A	14719	58923	39720	62118
UCP2	730406	607896	1292032	929678
POLR3B	10070	25233	14953	33109
ANAPC5	423120	306133	439400	404830
COL4A1	0	68	4	22
COL4A2	46	122	65	104

DHRS2	46656	14487	35513	29460
GALC	4532	18655	3607	9727
GAN	862	3822	488	1867
MPRIIP	17813	42913	24447	47359
PIGN	6987	33946	6636	20195
PNKP	195216	125001	188235	141200
WDR45	117655	92500	75813	79857
PLP1	0	33	0	13
MECP2	702270	510681	578258	481242

上の表で明らかなように、平均 100 万リードを超えるものから、8 リードしか読まれていない遺伝子までかなり幅が広がっており、いくらシーケンスをしても、リンパ芽球を用いた RNA 解析では解析不能の遺伝子 (PLB1、COL4A1、COL4A2、PLP1、GALNT13) が存在することが分かった。この問題を解決するためには、病気の場となる組織の RNA を得ることが重要だと考えられた。

(2) エクソーム解析で用いられる Nord script を用いて (引用文献 1)、エクソン毎のリード数の変化を基にしたエクソンスキップ検出の自動化を試みた。しかし、遺伝子毎に加えて、同じ遺伝子内でのエクソン毎のリードの厚みのばらつきも大きく (図 2)、標準偏差が大きい結果となったため (図 3)、エクソンスキップ検出の自動化は困難であった。

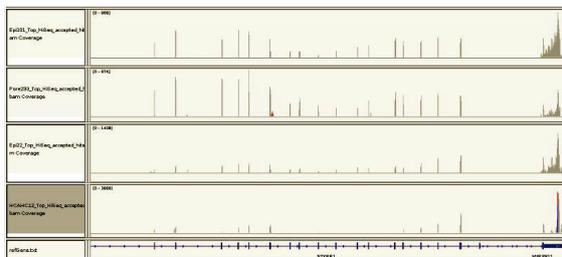


図 2 : 4 検体における STXPB1 遺伝子のリードの厚みを示す。各検体においても、また遺伝子内の各エクソンによっても大きくばらついている。

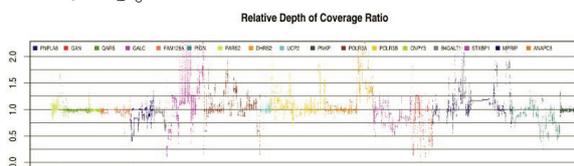


図 3 : Nord script を用いたエクソン毎のリードの厚みを用いたスプライス異常の検出の試み。ばらつきが大きく、この方法によるスプライス異常の自動検出は困難であった。

(3) 特定の遺伝子変異がもたらすスプライス異常を網羅的に解析するために、逆転写

PCR 産物にアダプターを付加して、MiSeq でシーケンスした。Novoalign でアライメントして、異常スプライス産物の同定およびその発現量の定量化を行った (図 4)。この解析によって、イントロンの変異によって 5 塩基の挿入と早期終止コドンの生成が起こり、異常転写産物がナンセンス変異依存性 mRNA 分解を受けることが明らかとなった。

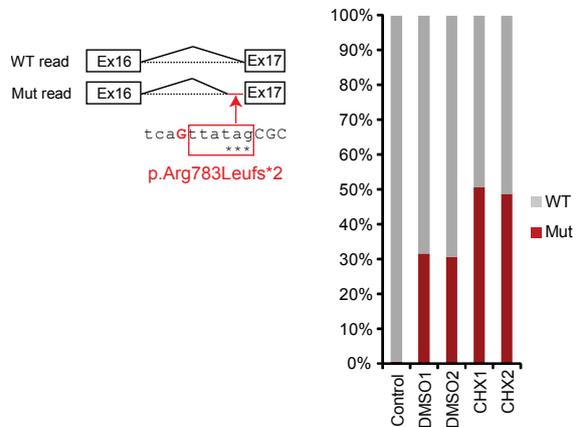


図 4 : 逆転写 PCR 産物を MiSeq でシーケンスして、Novoalign でゲノムにアライメントして、異常の確認 (左図) およびリード数のカウントによるナンセンス変異依存性 mRNA 分解の確認が可能であった。

<引用文献>

① Nord AS, Lee M, King MC, Walsh T. Accurate and exact CNV identification from targeted high-throughput sequence data. BMC Genomics. 2011 12;12:184.

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 1 件)

1. Kobayashi Y, Tohyama J, Kato M, Akasaka N, Magara S, Kawashima H, Ohashi T, Shiraishi H, Nakashima M, Saitsu H, Matsumoto N. High prevalence of genetic alterations in early-onset epileptic encephalopathies associated with infantile movement disorders. Brain and Development (査読あり) 38(3):285-92. DOI: 10.1016/j.braindev.2015.09.011.

[学会発表] (計 0 件)

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称 :
発明者 :
権利者 :

種類：
番号：
出願年月日：
国内外の別：

○取得状況（計 0 件）

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

才津 浩智 (SAITSU HIROTOMO)
浜松医科大学・医学部・教授
研究者番号：40402838

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：