



RNA sequencing and target long-read sequencing reveal an intronic transposon insertion causing aberrant splicing

メタデータ	言語: Japanese 出版者: 浜松医科大学 公開日: 2024-10-15 キーワード (Ja): キーワード (En): 作成者: 川上, 領太 メールアドレス: 所属:
URL	http://hdl.handle.net/10271/0002000246

博士（医学）川上 領太

論文題目

RNA sequencing and target long-read sequencing reveal an intronic transposon insertion causing aberrant splicing

（RNA シーケンスとターゲットロングリードシーケンスによるスプライシング異常の原因となるイントロンでのトランスポゾン挿入の同定）

論文の内容の要旨

〔はじめに〕

エクソームシーケンス (ES) やゲノムシーケンス (GS) はショートリードシーケンス (SRS) を利用しており、一塩基置換や小さい挿入欠失バリエントを正確に検出することが可能である。しかし、遺伝性疾患が疑われる症例の半数以上は ES や GS でも原因となる遺伝子異常が同定されない。その原因の一つとして、アミノ酸の置換を伴わない同義置換バリエントやアミノ酸に翻訳されない非翻訳領域のバリエントの解釈が難しいことが挙げられる。RNA シーケンス (RNA-seq) は遺伝子発現やスプライシングの変化を検出することが可能であるため、バリエントの解釈に有用である。また、SRS で得られるシーケンスは 150 塩基対 (bp) と短く、リピート配列の検出が困難である。リピート配列を含む構造変異の検出においては、数 kbp 以上の長いシーケンスが得られるロングリードシーケンス (LRS) が有用である。近年、ナノポアシーケンサーを用いた *adaptive sampling* によってターゲット LRS をより安価に行うことが可能となった。本研究では、常染色体潜性遺伝病が疑われたが ES、GS で 1 アレルの異常しか検出できなかった症例に対し、RNA-seq と LRS を実施した。

〔患者ならびに方法〕

患者は日本人の 3 歳女兒で、出生時から特徴的顔貌を指摘され、頭部 MRI で脳梁低形成、単純脳回、小脳虫部低形成が認められた。月齢 6 か月時に West 症候群と診断され、精神運動発達遅滞を指摘された。月齢 10 か月で摂食障害のため経管栄養が導入された。G 分染法による染色体検査で染色分体早期解離 (PCS) と多彩異数性モザイク (MVA) を指摘された。1 歳 3 か月時に腎芽腫、胎児型横紋筋肉腫を発症し、基礎疾患として PCS/MVA 症候群が疑われ、遺伝子解析を受けることとなった。本研究は浜松医科大学臨床研究倫理委員会の承認を得て行った（承認番号：20-207）。

患者末梢血由来 DNA を用いて ES、GS およびナノポアシーケンサーを用いた *adaptive sampling* によるターゲット LRS を実施した。同定されたバリエントに対する病的意義は複数の *in silico* 病的意義予測ツールで評価し、アレル頻度は公共データベースを参照した。病的バリエントの候補に対しては、患者及び両親の末梢血由来 DNA を用いてバリエントの確認を行った。また、患者及び母の尿由

来細胞（UDC）から抽出した RNA を用いて RNA-seq を実施した。スプライシング異常の確認のため、バリエーション前後の DNA 配列を発現ベクターに組み込み、培養細胞で発現させるミニ遺伝子スプライシングアッセイを実施した。また、タンパク質の発現量を、UDC を用いたイムノブロット法で評価した。

〔結果〕

ES では、PCS/MVA 症候群の常染色体潜性遺伝形式での責任遺伝子 *BUB1B* (NM_001211.6) に、母由来のヘテロ接合性バリエーション (c.1402-5A>G) を検出した。GS では父由来の病的バリエーションを検出できなかったが、RNA-seq では *BUB1B* のイントロン 2 に異常なリードを検出した。同部位をより詳細に検索するためナノポアシーケンサーでターゲット LRS を実施し、イントロン 2 に父由来である約 3.0 kbp のトランスポゾンの挿入を同定した。この挿入を含む 16 個のシーケンスリードに基づいてコンセンサス配列を作成したところ、この挿入は *Alu* 及び *SVA* エレメントで構成されていることが判明した。コンセンサス配列が挿入された新しい配列を参照配列として RNA-seq データを再アライメントしたところ、挿入内に 5'スプライス部位を、エクソン 3 に 3'スプライス部位を有する複数のシーケンスリードを認めたことから、挿入によりスプライシング異常が引き起こされていると予測した。逆転写 PCR によって RNA-seq で認められた複数のスプライシング異常を確認した。

イントロンバリエーションを有するアレルにおいて、それぞれの転写産物におけるスプライシング異常の割合を調べる目的で、遺伝子発現におけるアレルの不均衡を調べた。患者と母はヘテロ接合性バリエーション c.1046G>A を c.1402-5A>G と同一アレルに有していた。母の RNA-seq において c.1046G アレルが 46 リードであったのに対し、A アレルは 5 リードしか認めなかった。このアレル不均衡はミニ遺伝子スプライシングアッセイでも確認でき、c.1402-5G を有するアレルの転写産物のほぼ全てに 4 bp の挿入を受けるスプライシング異常が起こり、ナンセンス変異依存 mRNA 分解機構（NMD）を受けると考えられた。一方、患者では G アレルが 34 リード、A アレルは 12 リードであった。これは c.1046G アレルが 3.0 kbp の挿入と同一アレルにあるため、NMD によりある程度 G アレルが減少したことで、相対的に A アレルの割合が増加したと考えられる。更に、UDC を用いたイムノブロット法において、患者では母やコントロールと比較して *BUB1B* タンパク質の発現が顕著に減少していることを確認した。

〔考察〕

ES では母由来の c.1402-5A>G バリエーションを、LRS では父由来の 3.0 kbp の挿入を検出した。ミニ遺伝子スプライシングアッセイとアレル不均衡の解析により c.1402-5G を有するアレルがほぼヌルであることが示された。また、RNA-seq の再アライメントと母と患者のアレル不均衡の比較により 3.0 kbp の挿入がスプライシング異常を引き起こし、NMD を受けることが示された。更に、患者では

BUB1B の発現が著明に減少していることが示された。以上より c.1402-5A>G バリエーションと 3.0 kbp 挿入の両者が BUB1B の機能喪失を来し、PCS/MVA 症候群の原因となったことが示された。

トランスポゾン挿入を SRS で検出することは困難である。これはトランスポゾンが数百 bp から数 kbp と大きく、またヒトゲノム中にほぼ同一の配列を持つ多くのトランスポゾンが存在するためである。一方 LRS は 10 kbp 以上のリード長が得られることから、LRS はトランスポゾン挿入の検出において優れたツールであると言える。

LRS で挿入配列が得られたことは非常に重要である。PCR で切断点の確認を試みたが、5' 切断点は増幅させることができなかった。また、得られた挿入配列を用いることで RNA-seq の再アライメントが可能となり、挿入とスプライシング異常の因果関係を明らかにできた。

〔結論〕

RNA-seq や GS に LRS を加えたマルチオミクス研究は、非コーディング領域における病原性バリエーションの検出率を向上させる大きな可能性を有している